

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2005 年 9 月 9 日 (09.09.2005)

PCT

(10) 国際公開番号
WO 2005/083616 A1(51) 国際特許分類⁷: G06F 19/00,
G01N 33/48, 33/566, 33/68

(21) 国際出願番号: PCT/JP2005/003558

(22) 国際出願日: 2005 年 2 月 24 日 (24.02.2005)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願2004-048767 2004 年 2 月 24 日 (24.02.2004) JP

(71) 出願人 (米国を除く全ての指定国について): 株式会社インシリコサイエンス (IN-SILICO SCIENCES, INC.) [JP/JP]; 〒1450065 東京都大田区東雪谷二丁目 15 番 9 号 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 梅山 秀明

(UMEYAMA, Hideaki) [JP/JP]; 〒2790011 千葉県浦安市美浜 1-7-1002 Chiba (JP). 渡邊 佳晃 (WATANABE, Yoshiaki) [JP/JP]; 〒4120026 静岡県御殿場市東田中 15125-D Shizuoka (JP). 荒井 亮一 (ARAI, Ryoichi) [JP/JP]; 〒3530003 埼玉県志木市下宗岡 4-31-45 Saitama (JP).

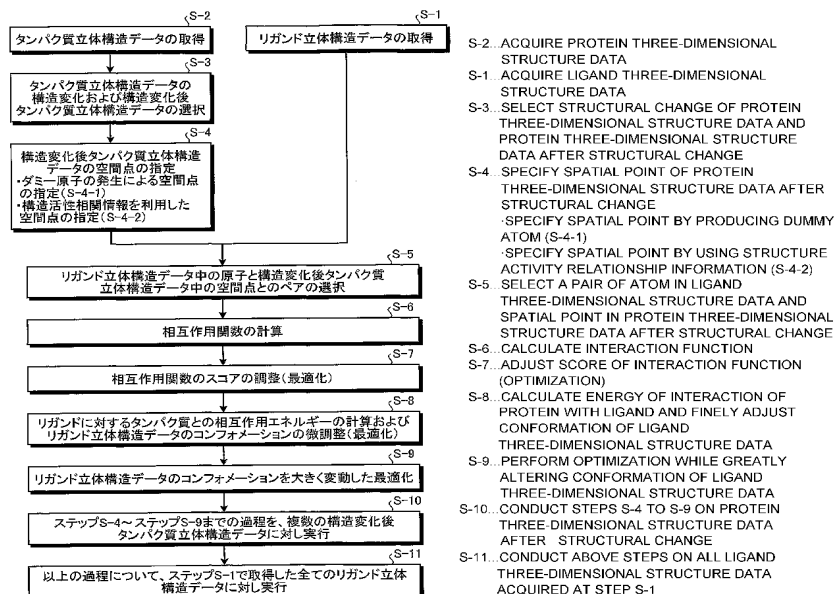
(74) 代理人: 酒井 宏明 (SAKAI, Hiroaki); 〒1000013 東京都千代田区霞が関三丁目 2 番 6 号 東京倶楽部ビルディング 酒井国際特許事務所 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[続葉有]

(54) Title: LIGAND SEARCHING DEVICE, LIGAND SEARCHING METHOD, PROGRAM, AND RECORDING MEDIUM

(54) 発明の名称: リガンド探索装置、リガンド探索方法、プログラム、および記録媒体



(57) Abstract: A ligand searching device for researching analysis of a receptor-ligand bond. The receptor can be an induced-fit receptor. A ligand searching method, a program, and a recording medium are also disclosed. Analysis calculation of the fundamental vibration of a receptor is conducted, and the fluctuation of the main chain dihedral angle of the receptor in a steady state is calculated. Molecular dynamics calculation is conducted while restraint based on the fluctuations is laid on each atom. Thus, the dynamic structure of a receptor is predicted with higher accuracy. Further, by using the dynamic structure determined by molecular dynamics calculation and an interaction function, the receptor/ligand bond applicable to an induced-fit receptor is predicted with high accuracy.

[続葉有]

WO 2005/083616 A1



(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告書
- 請求の範囲の補正の期限前の公開であり、補正書受領の際には再公開される。

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(57) 要約: 本発明は、誘導適合型受容体も含めた受容体／リガンド結合の解析を研究するためのリガンド探索装置、リガンド探索方法、プログラム、および記録媒体を提供することを目的とする。本発明は、受容体の基準振動解析計算をまず行い、定常状態の受容体の主鎖二面角の揺らぎを計算する。そして、その揺らぎを基にした拘束を各原子にかけながら分子動力学計算を行うことでより精度の良い受容体の動的構造を予測する。また、本発明は、分子動力学計算より得た動的構造及び相互作用関数を用いることで、誘導適合型受容体にも応用できる受容体／リガンド結合を精度良く予測する。

明 細 書

リガンド探索装置、リガンド探索方法、プログラム、および記録媒体

5 技術分野

本発明は、タンパク質の立体構造座標を用いたリガンド探索装置、リガンド探索方法、プログラム、および記録媒体に関し、特に、タンパク質の立体構造座標が既知の場合、相互作用すると考えられるリガンドを予測するリガンド探索装置、リガンド探索方法、プログラム、および記録媒体に関するものである。

10

背景技術

酵素や受容体等の生体機能を維持するために必要なタンパク質には、基質特異性と呼ばれる性質があり、活性部位が基質分子構造の細部にわたり常に一致しているLock&Key型と、基質が無いときには活性部位が不活性なランダムな状態にあり、基質が来るとこれを取り込むために活性部位が活性な状態に変化するInduced-Fit（誘導結合）型がある。誘導適合型とは、リガンドと結合する際にリガンド結合部位の立体構造が変化しリガンドを取り込むことが可能になる受容体をいう。

タンパク質の立体構造を用いたリガンド分子探索のための計算化学的手法としては、まず、AutoDock（「Morris, G. M. Goodsell, D. S. Halliday, R. S. Huey, R. Hart, W. E. Belew, R. K. Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J. Comput. Chem. 19: 1639-1662; Goodsell, D. S. Morris, G. M. Olson, A.

J. (1996) Automated docking of flexible ligands: applications of AutoDock. J. Mol. Recognit. 9: 1-5], DOCK (Ewing, T. J. Makino, S. Skillman, A. G. Kuntz I. D. (2001) DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. J. Comput. Aided Mol. Des. 15: 411-28]), FlexX (Rarey, M, Kramer, B, Lengauer, T, Klebe, G. (1996) A fast flexible docking method using an incremental construction algorithm. J. Mol. Biol. 261: 470-89; Rarey, M. Wefing, S. Lengauer, T. (1996) Placement of medium-sized molecular fragments into active sites of proteins. J. Comput. Aided Mol. Des. 10: 41-54]), GOLD (Jones, G. Willett, P. Glen, R. C. (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J. Mol. Biol. 245: 43-53; Jones, G. Willett, P. Glen, R. C. Leach, A. R. Taylor, R. (1997) Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. 267: 727-48]), ADAM&EVE (Mizutani,

M. Y. Itai, A. (2004) Efficient method for high-throughput virtual screening based on flexible docking: discovery of novel acetylcholinesterase inhibitors. J. Med. Chem. 47: 4818-4828」といった3次元化合物データベースサーチ (Virtual Screening) が知られている。これらは高速ドッキング・スタディーとも呼ばれ、大規模な化合物ライブラリ・サーチが可能である。しかし、本手法では評価に粗い近似を用いるため、結合配座や結合エネルギーの予測能は低い。さらにタンパク質とリガンドとの結合に大変重要な「誘導結合」に対応する計算式パラメータを充分に取り込んでいないので、たとえあったとしても、乱数を発生させ受容体の側鎖を動かす程度であり、計算結果の精度は充分なものとはいえない。

タンパク質とリガンドとの結合に重要な「誘導結合」をシミュレーションする方法としては、MD (分子動力学計算) やMM (分子力学計算)、MC (モンテカルロ法) が知られている。これらの方法は比較的精度良く、結合配座や結合エネルギーの予測が可能である。ここで、分子動力学法 (MD) と呼ばれる手法に関しては、ある分子を構成する各原子において、古典力学に基づく運動方程式を逐次的に解くことにより、その分子の動的構造を計算する方法であり、タンパク質の動的挙動を高精度でシミュレーションすることが可能である。しかし、計算に時間を要するため、多数の分子を扱うことは困難であり必ずしも有用な手法とはなっていない。さらに、従来法では該当タンパク質に対して分子動力学計算を行うと、タンパク質立体構造はX線やNMR等で解析された座標から大きくズレる。こうしたズレはタンパク質の動的挙動の物理化学的描写を含んでいるが、NMR等で示される動的挙動の実験的な結果と矛盾する挙動となる場合があり、必ずしも精度の高いシミュレーションとならないことが多い。

このように、従来の *in silico* screening 関連では、タンパク質とリガンドとの結合に大変重要な「誘導結合」に対応する計算式パラメー

タを充分に取り込んでいないので、計算結果の精度は充分なものとはいえない。

一方、分子シミュレーションでは、上記の誘導結合を表現し解析することは可能であるが、高精度の結果を得るためには相当の時間を必要とする。多くの結果は、初期構造座標に依存してしまう。

- 5 本発明者等は、任意のタンパク質の立体構造が与えられたとき、該当タンパク質に結合するリガンドを探索する方法について検討した。上述したように、現在流通している受容体・リガンド結合解析ソフトには、リガンドのフレキシビリティを考慮しているものはあるが、受容体側のフレキシビリティを考慮しているものはほとんどない。また、たとえあったとしても、乱数を発生させ受容体の側鎖を動かす程度であり、L o c k & K e y型の受容体に対応しているものばかりであつた。そこで、I n d u c e d - F i t型の受容体を対象にした受容体・リガンド結合解析ソフトを開発することにした。
- 10

- 本発明が解決しようとする課題は、農薬や医薬品等の開発に特に重要な鍵となる、該当タンパク質に結合するリガンドを、従来法に比べてはるかに効率的に精度よく探索するリガンド探索装置、リガンド探索方法、プログラム、および記録媒体を提供することである。また、リガンド分子の多様な改変や受容体等のタンパク質の改変を迅速かつ効率的に行うリガンド探索装置、リガンド探索方法、プログラム、および記録媒体を提供することである。さらに、本発明により、リガンドータンパク質間の相互作用様式を解明し、それら相互作用の認識機構を明確化することで、疾病の原因を特定したり、関連する薬物の開発を促進したりすること等を目的とする。
- 15
- 20

発明の開示

- 本発明者等は、任意のタンパク質立体構造が与えられたとき、該当タンパク質に結合するリガンドを探索する方法について検討を重ねた結果、下記のリガンド探索装置、リガンド探索方法、コンピュータプログラム、および記録媒体を開発した。
- 25

ここで、分子動力学法（MD）と呼ばれる手法があり、これはある分子を構成する各原子において、古典力学に基づく運動方程式を逐次的に解くことにより、その分子の動的構造を計算する方法である。つまり、これはある分子を構成する各原子における古典力学を土台とした動的挙動を計算する方法である。従って、

5 この手法をうまく取り込むことができれば、リガンドを取り込んでいない状態の Induced-Fit 型受容体を初期状態に選んでも、受容体・リガンド結合を再現できると考えた。MD 計算は古典力学を土台にしているため、各原子にある程度の拘束をかける必要がある。そこで、まず初めに受容体の基準振動解析を行い受容体の主鎖二面角揺らぎを計算し、この主鎖二面角揺らぎに基づいて各原子に拘束をかけて MD を計算する手法を開発した。具体的には、基準振動解析計算をまず行い、定常状態の主鎖二面角の揺らぎを計算する。そして、その揺らぎを基にした拘束を各原子にかけながら分子動力学計算を行うことで、受容体の動的構造をより精度良く予測する。また、分子動力学計算より得た動的構造及び相互作用関数を用いることで、誘導適合型受容体にも応用できる受容体／リガンド

10 結合を精度良く予測する。つまり、本発明は、より真に近い受容体／リガンド結合を精度良く予測する。従って、本発明は、医薬分子の設計に極めて有用である。

このような目的を達成するために、本発明にかかるリガンド探索装置は、単数または複数鎖のタンパク質の座標データが与えられた場合に、当該タンパク質と結合するリガンドを探索するリガンド探索装置において、上記タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択する構造変化後タンパク質座標データ選択手段と、上記構造変化後タンパク質座標データ選択手段にて選択された上記構造変化後タンパク質座標データから、上記リガンドと重ね合わせを行う空間点を指定する空間点指定手段と、上記空間点指定手段にて指定された上記空間点と、上記リガンドのリガンド座標データとを用いて、上記タンパク質と上記リガンドとが結合した場合の相互作用関数を計算する相互作用関

20

25

数計算手段と、上記相互作用関数計算手段により計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドを評価するリガンド評価手段と、を備えたことを特徴とする。

また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記相互作用関数計算手段は、以下の数式1に示す $Sscore(i, j)$ を用いて上記相互作用関数を計算することを特徴とする。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp\left\{-\left(d_{ij}^s - d_{ij}^c\right)^2\right\} - \beta \right] / \frac{\left(d_{ij}^s + d_{ij}^c\right)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases} \quad \dots \text{ (数式1)}$$

(ここで、 d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。また、 α は、該当タンパク質中の空間点群と化合物が完全に重なりあった場合に $Sscore(i, j)$ を最大値とするための定数である。また、 β は重なりと定義できる限界値を与えるための定数である。)

また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記相互作用関数計算手段は、上記相互作用関数のスコアが最大になるように最適化する相互作用関数最適化手段、をさらに備えたことを特徴とする。

また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記相互作用関数計算手段は、上記相互作用関数最適化手段により上記相互作用関数を最適化した後に、重ねあわせた上記リガンドに対して、上記タンパク質との相互作用エネルギーを計算し、当該相互作用エネルギーについて

リガンド立体構造データのコンフォメーションを微調整しながら最適化する相互作用エネルギー最適化手段、をさらに備えたことを特徴とする。

また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記リガンド評価手段は、上記相互作用エネルギー最適化手段により最適化した後に、上記リガンド立体構造データのコンフォメーションを大きく変動させた後、再度、上記相互作用関数計算手段を実行し、上記相互作用関数計算手段により計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドの再評価を行う再評価手段、をさらに備えたことを特徴とする。

また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記構造変化後タンパク質座標データ選択手段は、上記誘導適合パラメータおよび／または上記構造変化後タンパク質座標データを算出するときに、上記タンパク質座標データに対し基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、当該ゆらぎの大きさを拘束条件として、分子動力学計算を行うこと、を特徴とする。

また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記構造変化後タンパク質座標データ選択手段は、基準振動計算より主鎖原子の2面角のゆらぎの値を算出し、当該ゆらぎ値を以下の数式2または数式3に示す分子動力学計算における力の定数Kとすることにより、分子動力学計算を行うこと、を特徴とする。

$$E_{rot} = K_{rot} (\phi - \phi_0)^2 \quad \dots \text{(数式2)}$$

(E_{rot} はタンパク質の立体構造中において主鎖原子の2面角のエネルギーを示す。 ϕ は主鎖原子の2面角である。 ϕ_0 は主鎖原子の2面角の標準値である。

ここで、 K_{rot} の値が大きい場合は、 ϕ は ϕ_0 に拘束される。)

$$E_{pos} = K_{pos} (r - r_0)^2 \quad \dots \text{(数式3)}$$

(E p o s はタンパク質の立体構造中において主鎖原子の位置のエネルギーを示す。r は主鎖原子の座標である。r 0 は主鎖原子の座標の標準値である。ここで、K p o s の値が大きい場合は、r はr 0 に拘束される。)

- 5 また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記相互作用関数計算手段は、上記相互作用関数として、相互作用エネルギー関数にタンパク質の動的性質を表現する動的性質関数を「弾性エネルギー」として加えて用いること、を特徴とする。

- 10 また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記相互作用関数計算手段は、「弾性エネルギー」として、タンパク質の局所的な柔らかさを考慮し、以下の数式4に示す関数「U衝突」を適用すること、を特徴とする。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

- 15 $\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$

... (数式4)

- (Mは衝突を不許可とする活性部位の原子の数であり、Nはリガンドの原子の数である。活性部位における動的挙動の少ない側鎖原子及び主鎖原子のみのi番目の原子とリガンドのj番目の原子との原子間距離Rが衝突距離「R衝突(i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。)
- 20

- また、つぎの発明にかかるリガンド探索装置は、上記に記載のリガンド探索装置において、上記相互作用関数計算手段は、上記相互作用関数として、相互作用エネルギー関数に、当該タンパク質の基準振動解析結果または二次構造判定結果をタンパク質の動的性質を表現する動的性質関数として加えて用いること、を特徴とする。
- 25

また、本発明はリガンド探索方法に関するものであり、本発明にかかるリガンド探索方法は、単数または複数鎖のタンパク質の座標データが与えられた場合に、

- 当該タンパク質と結合するリガンドを探索するリガンド探索方法において、上記タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択する構造変化後タンパク質座標データ選択ステップと、上記構造変化後タンパク質座標データ選択ステップにて選択された上記構造変化後タンパク質座標データから、上記リガンドと重ね合わせを行う空間点を指定する空間点指定ステップと、上記空間点指定ステップにて指定された上記空間点と、上記リガンドのリガンド座標データとを用いて、上記タンパク質と上記リガンドとが結合した場合の相互作用関数を計算する相互作用関数計算ステップと、上記相互作用関数計算ステップにより計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドを評価するリガンド評価ステップと、を含むことを特徴とする。

また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記相互作用関数計算ステップは、以下の数式1に示す $Sscore(i, j)$ を用いて上記相互作用関数を計算することを特徴とする。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp\left\{-\left(d_{ij}^s - d_{ij}^c\right)^2\right\} - \beta \right] / \frac{\left(d_{ij}^s + d_{ij}^c\right)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases} \quad \dots (数式1)$$

- (ここで、 d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。また、 α は、該当タンパク質中の空間点群と化合物が完全に重なりあった場合に $Sscore(i, j)$ を最大値とするための定数である。また、 β は重なりと定義できる限界値を与えるための定数である。)

また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記相互作用関数計算ステップは、上記相互作用関数のスコアが最大になるように最適化する相互作用関数最適化ステップ、をさらに含むことを特徴とする。

- 5 また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記相互作用関数計算ステップは、上記相互作用関数最適化ステップにより上記相互作用関数を最適化した後に、重ねあわせた上記リガンドに対して、上記タンパク質との相互作用エネルギーを計算し、当該相互作用エネルギーについてリガンド立体構造データのコンフォメーションを微調整しながら最適化
10 する相互作用エネルギー最適化ステップ、をさらに含むことを特徴とする。

- また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記リガンド評価ステップは、上記相互作用エネルギー最適化ステップにより最適化した後に、上記リガンド立体構造データのコンフォメーションを大きく変動させた後、再度、上記相互作用関数計算ステップを実行し、上記相
15 互作用関数計算ステップにより計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドの再評価を行う再評価ステップ、をさらに含むことを特徴とする。

- また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記構造変化後タンパク質座標データ選択ステップは、上記誘導適
20 合パラメータおよび／または上記構造変化後タンパク質座標データを算出するときに、上記タンパク質座標データに対し基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、当該ゆらぎの大きさを拘束条件として、分子動力学計算を行うこと、を特徴とする。

- また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記構造変化後タンパク質座標データ選択ステップは、基準振動計
25 算より主鎖原子の2面角のゆらぎの値を算出し、当該ゆらぎ値を以下の数式2または数式3に示す分子動力学計算における力の定数Kとすることにより、分子動

力学計算を行うこと、を特徴とする。

$$E_{rot} = K_{rot} (\phi - \phi_0)^2 \quad \dots \text{(数式2)}$$

- 5 (E_{rot}はタンパク質の立体構造中において主鎖原子の2面角のエネルギーを示す。 ϕ は主鎖原子の2面角である。 ϕ_0 は主鎖原子の2面角の標準値である。ここで、K_{rot}の値が大きい場合は、 ϕ は ϕ_0 に拘束される。)

$$E_{pos} = K_{pos} (r - r_0)^2 \quad \dots \text{(数式3)}$$

10

(E_{pos}はタンパク質の立体構造中において主鎖原子の位置のエネルギーを示す。 r は主鎖原子の座標である。 r_0 は主鎖原子の座標の標準値である。ここで、K_{pos}の値が大きい場合は、 r は r_0 に拘束される。)

- また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記相互作用関数計算ステップは、上記相互作用関数として、相互作用エネルギー関数にタンパク質の動的性質を表現する動的性質関数を「弾性エネルギー」として加えて用いること、を特徴とする。

- また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記相互作用関数計算ステップは、「弾性エネルギー」として、タンパク質の局所的な柔らかさを考慮し、以下の数式4に示す関数「U衝突」を適応すること、を特徴とする。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

25

... (数式4)

(Mは衝突を不許可とする活性部位の原子の数であり、Nはリガンドの原子の数である。活性部位における動的挙動の少ない側鎖原子及び主鎖原子のみのi番

目の原子とリガンドの j 番目の原子との原子間距離 R が衝突距離「 R 衝突 (i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。)

また、つぎの発明にかかるリガンド探索方法は、上記に記載のリガンド探索方法において、上記相互作用関数計算ステップは、上記相互作用関数として、相互作用エネルギー関数に、当該タンパク質の基準振動解析結果または二次構造判定結果をタンパク質の動的性質を表現する動的性質関数として加えて用いること、を特徴とする。

また、本発明はプログラムに関するものであり、本発明にかかるリガンド探索方法をコンピュータに実行させることを特徴とするプログラムは、単数または複数鎖のタンパク質の座標データが与えられた場合に、当該タンパク質と結合するリガンドを探索するリガンド探索方法をコンピュータに実行させるプログラムにおいて、上記タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択する構造変化後タンパク質座標データ選択ステップと、上記構造変化後タンパク質座標データ選択ステップにて選択された上記構造変化後タンパク質座標データから、上記リガンドと重ね合わせを行う空間点を指定する空間点指定ステップと、上記空間点指定ステップにて指定された上記空間点と、上記リガンドのリガンド座標データとを用いて、上記タンパク質と上記リガンドとが結合した場合の相互作用関数を計算する相互作用関数計算ステップと、上記相互作用関数計算ステップにより計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドを評価するリガンド評価ステップと、を含むことを特徴とする。

また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記相互作用関数計算ステップは、以下の数式1に示す $Sscore(i, j)$ を用いて上記相互作用関数を計算することを特徴とする。

$$Sscore(i, j) = \sum_j^A \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp\{-(d_{ij}^s - d_{ij}^c)^2\} - \beta \right] / \frac{(d_{ij}^s + d_{ij}^c)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases}$$

5

・・・ (数式1)

- (ここで、 d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。
 10 また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。また、 α は、該
 当タンパク質中の空間点群と化合物が完全に重なりあった場合に $Sscore(i, j)$ を最大値とするための定数である。また、 β は重なりと定義できる限界
 値を与えるための定数である。)

- また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、
 15 上記相互作用関数計算ステップは、上記相互作用関数のスコアが最大になるよう
 に最適化する相互作用関数最適化ステップ、をさらに含むことを特徴とする。

- また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、
 上記相互作用関数計算ステップは、上記相互作用関数最適化ステップにより上記
 相互作用関数を最適化した後に、重ねあわせた上記リガンドに対して、上記タン
 20 パク質との相互作用エネルギーを計算し、当該相互作用エネルギーについてリガ
 ンド立体構造データのコンフォメーションを微調整しながら最適化する相互作用
 エネルギー最適化ステップ、をさらに含むことを特徴とする。

- また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、
 上記リガンド評価ステップは、上記相互作用エネルギー最適化ステップにより最
 25 適化した後に、上記リガンド立体構造データのコンフォメーションを大きく変動
 させた後、再度、上記相互作用関数計算ステップを実行し、上記相互作用関数計
 算ステップにより計算された上記相互作用関数に基づいて当該タンパク質と結合

する上記リガンドの再評価を行う再評価ステップ、をさらに含むことを特徴とする。

また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記構造変化後タンパク質座標データ選択ステップは、上記誘導適合パラメータ
5 および／または上記構造変化後タンパク質座標データを算出するときに、上記タンパク質座標データに対し基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、当該ゆらぎの大きさを拘束条件として、分子動力学計算を行うこと、を特徴とする。

また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、
10 上記構造変化後タンパク質座標データ選択ステップは、基準振動計算より主鎖原子の2面角のゆらぎの値を算出し、当該ゆらぎ値を以下の数式2または数式3に示す分子動力学計算における力の定数Kとすることにより、分子動力学計算を行うこと、を特徴とする。

$$15 \quad E_{rot} = K_{rot} (\phi - \phi_0)^2 \quad \dots \text{(数式2)}$$

(E_{rot} はタンパク質の立体構造中において主鎖原子の2面角のエネルギーを示す。 ϕ は主鎖原子の2面角である。 ϕ_0 は主鎖原子の2面角の標準値である。ここで、 K_{rot} の値が大きい場合は、 ϕ は ϕ_0 に拘束される。)

20

$$E_{pos} = K_{pos} (r - r_0)^2 \quad \dots \text{(数式3)}$$

(E_{pos} はタンパク質の立体構造中において主鎖原子の位置のエネルギーを示す。 r は主鎖原子の座標である。 r_0 は主鎖原子の座標の標準値である。ここで、 K_{pos} の値が大きい場合は、 r は r_0 に拘束される。)

25

また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記相互作用関数計算ステップは、上記相互作用関数として、相互作用エネルギー

一関数にタンパク質の動的性質を表現する動的性質関数を「弾性エネルギー」として加えて用いること、を特徴とする。

また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記相互作用関数計算ステップは、「弾性エネルギー」として、タンパク質の局
5 所的な柔らかさを考慮し、以下の数式4に示す関数「U衝突」を適応すること、を特徴とする。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

10 ・・・ (数式4)

(Mは衝突を不許可とする活性部位の原子の数であり、Nはリガンドの原子の数である。活性部位における動的挙動の少ない側鎖原子及び主鎖原子のみのi番目の原子とリガンドのj番目の原子との原子間距離Rが衝突距離「R衝突(i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。)

15 また、つぎの発明にかかるプログラムは、上記に記載のプログラムにおいて、上記相互作用関数計算ステップは、上記相互作用関数として、相互作用エネルギー関数に、当該タンパク質の基準振動解析結果または二次構造判定結果をタンパク質の動的性質を表現する動的性質関数として加えて用いること、を特徴とする。

また、本発明は記録媒体に関するものであり、本発明にかかる記録媒体は、
20 上記に記載されたプログラムを記録したことを特徴とする。

図面の簡単な説明

第1図は、本発明の基本原理を示す概念図であり、第2図は、s p²軌道原子におけるダミー水素原子を示す図であり、第3図は、金属原子におけるダミー原子を示す図であり、第4図は、構造活性相関(SAR)情報をもとに活性部位内にリガンドを入れるための初期座標(B)を示す図であり、第5図は、構造活性相関(SAR)情報をもとに活性部位内にリガンドを入れるための初期座標(B
25

）を示す図であり、第6図は、構造活性相関（SAR）情報をもとに活性部位内にリガンドを入れるための初期座標（B）を示す図であり、第7図は、構造活性相関（SAR）情報をもとに活性部位内にリガンドを入れるための初期座標（B）を示す図であり、第8図は、構造活性相関（SAR）情報をもとに活性部位内にリガンドを入れるための初期座標（B）を示す図であり、第9図は、水素結合角の定義を示す図であり、第10図は、スタッキングにおける角度の定義を示す図であり、第11図は、1 LUDのMODEL 1における基準振動解析の結果を示す図であり、第12図は、MD及びクラスタリングのパラメータとスコアとの比較の結果を示す図であり、第13図は、クラスタリング定数を固定した時のMDでの二面角拘束の最大値および最小値の分布を示す図であり、第14図は、拘束パラメータを示す図であり、第15図は、クラスタリングのパラメータを固定した時の二面角拘束パラメータの分布を示す図であり、第16図は、クラスタリングのパラメータを固定した時の二面角拘束パラメータの分布を示す図であり、第17図は、クラスタリングのパラメータを固定した時の二面角拘束パラメータの分布を示す図であり、第18図は、クラスタリングのパラメータを固定した時の二面角拘束パラメータの分布を示す図であり、第19図は、1 LUD（MODEL 1）のMDの結果を示す図であり、第20図は、二面角拘束なしでMDを計算させた時のNMR構造との比較結果を示す図であり、第21図は、二面角拘束ありでMDを計算させた時のNMR構造との比較結果を示す図であり、第22図は、1 C B Qのアライメントを示した図であり、第23図は、1 C B Qの立体構造を示す図であり、第24図は、1 C B Qの立体構造を示す図であり、第25図は、1 C B QのX線構造とモデル構造の相違をrmsで表示した図であり、第26図は、1 C B QのX線構造とモデル構造の基準振動解析の結果を示す図であり、第27図は、1 C B QのX線構造とモデル構造の基準振動解析の結果を示す図であり、第28図は、1 C B QのX線構造とモデル構造のMD計算の結果を示す図であり、第29図は、1 J 9 Gのアライメントを示す図であり、第30図は、1 J 9 Gの立体構造を示す図であり、第31図は、1 J 9 Gの立体構造を示す図で

あり、第32図は、1J9GのX線構造とモデル構造の相違をrmsで表示した図であり、第33図は、1J9GのX線構造とモデル構造の基準振動解析の結果を示す図であり、第34図は、1J9GのX線構造とモデル構造の基準振動解析の結果を示す図であり、第35図は、1J9GのX線構造とモデル構造のMD計算の結果を示す図であり、第36図は、1MMBのアライメントを示す図であり、第37図は、1MMBの立体構造を示す図であり、第38図は、1MMBの立体構造を示す図であり、第39図は、1MMBのX線構造とモデル構造の相違をrmsで表示した図であり、第40図は、1J9GのX線構造とモデル構造の基準振動解析の結果を示す図であり、第41図は、1J9GのX線構造とモデル構造の基準振動解析の結果を示す図であり、第42図は、1MMBのX線構造とモデル構造のMD計算の結果を示す図であり、第43図は、ジヒドロ葉酸還元酵素の立体構造を示す図であり、第44図は、1BZF (MODEL18) の基準振動解析の結果を示す図であり、第45図は、1BZF (MODEL18) のMD計算の結果を示す図であり、第46図は、1LUD (MODEL4) より得られた構造活性相関情報を示す図であり、第47図は、1BZF (MODEL18) における活性部位・リガンド結合解析の結果を示す図であり、第48図は、1BZF (MODEL4) ・リガンド結合を示す図であり、第49図は、1BZF (MODEL4) ・リガンド結合を示す図であり、第50図は、1BZF (MODEL4) ・リガンド結合を示す図であり、第51図は、heat shock protein 90の立体構造を示す図であり、第52図は、1YERの基準振動解析の結果を示す図であり、第53図は、1YERのMD計算の結果を示す図であり、第54図は、1YETより得られた構造活性相関情報を示す図であり、第55図は、1YERにおける活性部位・リガンド結合解析の結果を示す図であり、第56図は、1YER・リガンド結合を示す図であり、第57図は、1YER・リガンド結合を示す図であり、第58図は、mitogen-activated protein kinaseの立体構造を示す図であり、第59図は、1A9Uの基準振動解析の結果を示す図であり、第60図は、1A9UのMD計

算の結果を示す図であり、第61図は、1OUKより得られた構造活性相関情報を示す図であり、第62図は、1A9Uにおける活性部位・リガンド結合解析の結果を示す図であり、第63図は、1A9U・リガンド結合を示す図であり、第64図は、1A9U・リガンド結合を示す図であり、第65図は、1A9U・リ
5 ガンド結合を示す図であり、第66図は、1AIXの立体構造を示す図であり、第67図は、in silico screeningの結果を示す図であり、第68図は、タンパク質／リガンド複合体構造を示す図であり、第69図は、タンパク質／リガンド複合体構造を示す図であり、第70図は、タンパク質／リガンド複合体構造を示す図であり、第71図は、タンパク質／リガンド複合体構造
10 を示す図であり、第72図は、タンパク質／リガンド複合体構造を示す図であり、第73図は、タンパク質／リガンド複合体構造を示す図であり、第74図は、タンパク質／リガンド複合体構造を示す図であり、第75図は、タンパク質／リガンド複合体構造を示す図であり、第76図は、SARSプロテアーゼの立体構造を示す図であり、第77図は、1UK3（B鎖）の基準振動解析の結果を示す図
15 であり、第78図は、1UK3（B鎖）のMD計算の結果を示す図であり、第79図は、1UK4（B鎖）より得られた構造活性相関情報を示す図であり、第80図は、1UK3（B鎖）におけるin silicoスクリーニングの結果を示す図であり、第81図は、1UK3と1UK4との比較結果を示す図であり、第82図は、in silicoスクリーニングの順位1を示す図であり、第8
20 3図は、in silicoスクリーニングの順位1を示す図であり、第84図は、1UK3（B鎖）より得られた構造活性相関情報を示す図であり、第85図は、1UK3（B鎖）と1UK4（B鎖）との比較結果を示す図であり、第86図は、SAR3カ所指定で実行したin silicoスクリーニングの結果を示す図
であり、第87図は、1UK3（B鎖）より得られた構造活性相関情報を示す図
25 であり、第88図は、SAR5カ所指定で実行したハイスループットスクリーニングの結果を示す図であり、第89図は、1UK3（B鎖）と1UK4（B鎖）との比較結果を示す図であり、第90図は、1UK3（B鎖）より得られた構造

活性相関情報を示す図であり、第91図は、リガンド原子タイプ指定変更で実行したハイスループットスクリーニングの結果を示す図であり、第92図は、1UK3 (B鎖) と1UK4 (B鎖) との比較結果を示す図であり、第93図は、1UK4 (B鎖) より得られた構造活性相関情報を示す図であり、第94図は、受容体を固定した状態で実行したハイスループットスクリーニングの結果を示す図であり、第95図は、1UK3と1UK4を重ね合わせたリガンドと計算結果のリガンドとの比較結果を示す図であり、第96図は、1AXJにおける二面角拘束MDパラメータの分布を示す図であり、第97図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第98図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第99図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第100図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第101図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第102図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第103図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第104図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第105図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第106図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第107図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第108図は、1LUD (MODEL 1) のMD計算の結果を示す図であり、第109図は、異なる条件で受容体／リガンドの結合結果を示す図であり、第110図は、異なる条件で受容体／リガンドの結合結果を示す図であり、第111図は、異なる条件で受容体／リガンドの結合結果を示す図であり、第112図は、1BZF用に変更した構造活性相関情報を示す図であり、第113図は、1BZF (MODEL 18) のリガンド結合解析の結果を示す図であり、第114図は、1BZF (MODEL 18) のリガンド結合解析の結果を示す図であり、第115図は、本実施形態における本システムのメイン処理の一例を示すフローチャートであり、第116図は、本発明が適用される本システムの構成の

一例を示すブロック図であり、第117図は、本発明が適用される本システムの相互作用関数計算部102cの構成の一例を示すブロック図であり、第118図は、本発明が適用される本システムのリガンド評価部102dの構成の一例を示すブロック図である。

5

発明を実施するための最良の形態

以下に、本発明にかかるリガンド探索装置、リガンド探索方法、プログラム、および記録媒体の実施の形態を図面に基づいて詳細に説明する。なお、この実施の形態により本発明が限定されるものではない。

10 本明細書において幾つかの用語を使用するが、特に明記しない限り、次の意味を有する。

「標的タンパク質」とは、立体構造の詳細がX線結晶解析やNMR解析、ホモロジーモデリング法により既に決定されており、リガンド探索の対象とするタンパク質を意味する。

15 「原子座標」とは、三次元空間上で立体構造を記述するものである。それは空間上のある点を原点とする互いに垂直な三方向の相対的な距離であり、タンパク質中に存在する水素原子を除く原子1つあたりに3個の数字からなるベクトル量である。

[本発明の基本原理]

20 ここでは、本発明の基本原理について、第1図を参照して説明する。第1図は、本発明の基本原理を示す概念図である。本発明は、単数または複数鎖のタンパク質立体構造が与えられた場合において、当該タンパク質の立体構造から誘導適合を反映したパラメータおよび構造変化した立体構造座標を例えば基準振動計算方法や分子動力学計算方法により予め算出し、当該パラメータおよび構造変化した
25 立体構造座標を用いて該当タンパク質と別の物質（リガンド）が結合した場合の相互作用関数を定義し、当該相互作用関数によって該当タンパク質と結合する物質（リガンド）を評価し、選定するリガンド探索装置、リガンド探索方法、コン

ピュータプログラム、および記録媒体に関する。

まず、本発明は、化合物データベースからリガンドを1つ選択し、そのリガンドの立体構造データを取得する（ステップS-1）。また、対象となるタンパク質の立体構造データを取得する（ステップS-2）。

- 5 つぎに、本発明は、当該タンパクの立体構造データに基づいて、誘導適合を反映する誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、複数の構造変化後タンパク質座標データを用意し、ランダムに1つの構造変化後タンパク質座標データを選択する（ステップS-3）。

- 10 つぎに、本発明は、リガンドと重ね合わせを行う、構造変化後タンパク質座標データ中の空間点を指定する（ステップS-4）。ここで、当該空間点は、例えば、以下に示す（1）や（2）などの方法で指定してもよい。

（1）ダミー原子の発生による空間点の指定（ステップS-4-1）

- 15 リガンドとタンパク質との相互作用における水素結合に着目し、タンパク質中の水素結合サイトを空間点として指定する。水素結合における重要事項は距離と角度である。つまり、角度を計算するためには、水素結合ドナー（以後、ドナー）に水素原子が必要になる。

そこで、本発明では、活性部位及びリガンドに水素原子が含まれていない場合、以下の規則によりダミー水素原子を発生させる。

- 20 1) sp^2 軌道原子を中心とする正三角形にダミー原子を発生させる（第2図）。すなわち、第2図に示すように、 sp^2 軌道原子の窒素原子（A）を中心とする正三角形の空いている位置にダミー水素原子（B）を発生させる。
- 2) sp^3 軌道原子では、水素結合を形成する距離にある場合、水素原子を共有するように回転できると考え、水素結合相互作用を計算するときには距離のみを考慮する。このため、 sp^3 軌道原子にはダミー原子を発生させない。

- 25 また、金属及び水では、活性部位・リガンド結合の仲介役となり得るので、相互作用する位置に以下のようにダミー原子を発生させる。

1) 鉄のような金属には、正八面体状にダミー原子を発生させる（第3図）。す

なわち、第3図に示すように、亜鉛（A）を中心とする正八面体の空いている位置にダミー原子（B）を発生させる。

2）水には、正四面体状にダミー原子を発生させる。

ただし、活性部位と相互作用している方向にはダミー原子を発生させなくてもよい。

（2）構造活性相関情報を利用した空間点の指定（ステップS-4-2）

また、本発明は、リガンドの構造活性相関（SAR）情報に着目し、以下の（A）～（D）の項目を入力情報にすることにより、空間点を指定する。

（A）SAR情報から得られた活性部位の原子（以後、「A原子」）。PDB形式に従う。

（B）「A原子」と相互作用するであろうリガンドの原子タイプ（以後、「Bタイプ」）。SYBYLのMOL2形式に従う。

（C）「A原子」と「Bタイプ」との相互作用の強さ（以後、「C強さ」）。

（D）「A原子」と「Bタイプ」との相互作用する距離（以後、「D距離」）（単位はÅ）。

なお、本発明では、上記（A）～（D）の入力情報に基づいて、タンパク質中の活性部位内におけるリガンドの初期座標を利用し、以下の1）～4）に示す規則により空間点を作成してもよい。

1）「A原子」がドナーまたは金属及び水の場合（SAR情報の活性部位側の指定が水素結合ドナー、金属原子の場合）には、ステップS-4-1で発生させたダミー原子の方向に対して「A原子」から「D距離」の位置及びその周囲を初期座標に選ぶ（第4図、第5図）。

2）「A原子」が sp^3 軌道原子の場合（SAR情報の活性部位側の指定が sp^3 軌道原子の場合）には、「A原子」から「D距離」の周囲を初期座標に選ぶ（第6図）。

3）「A原子」が水素結合アクセプター（以後、アクセプター）の場合（SAR情報の活性部位側の指定が水素結合アクセプターの場合）には、「A原子」の結

合延長上「D距離」の位置及びその周囲を初期座標に選ぶ（第7図）。

4) その他の場合（SAR情報の活性部位側の指定がその他の原子の場合）には、「A原子」を中心とする半径が「D距離」の球表面上の点を初期座標に選ぶ（第8図）。

5) ここで、上記1)～4)とは異なり、リガンドの初期座標を直接指定してもよい。

再び第1図に戻り、本発明は、ステップS-1において選択したリガンド座標データ中の原子と、ステップS-4で指定したタンパク質座標データ中の空間点とのペアを重複がないようにランダムに選択する（ステップS-5）。

10) つぎに、本発明は、以下の数式1に示す相互作用関数であるスコア $Sscore(i, j)$ を計算する（ステップS-6）。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp\left\{-\frac{(d_{ij}^s - d_{ij}^c)^2}{2}\right\} - \beta \right] / \frac{(d_{ij}^s + d_{ij}^c)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases} \quad \dots (数式1)$$

20) ここで、 d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。 α は、該当タンパク質中の空間点群と化合物が完全に重なりあった場合に $Sscore(i, j)$ を最大値とするための定数である。 β は重なりと定義できる限界値を与えるための定数である。

25) また、 α は1.5、 β は0.8とするのが好ましい。

つぎに、本発明は、ステップS-6により求めた相互作用関数のスコアが最大になるように調整（最適化）する（ステップS-7）。ここで、スコアを最大に

する手法としては、例えば、シミュレーティッドアニーリング法が挙げられる。
また、時間短縮にはステップS-5およびステップS-6を複数回（例えば10
000回）繰り返し、 $Sscore(i, j)$ が最大になるペアを探し、そのペ
ア情報をもとにリガンドを初期座標に重ね合わせる方法を適応することが好まし
5 い。

つぎに、本発明は、ステップS-7により相互作用関数を最適化したときに、
重ね合わせたリガンドに対して、タンパク質との相互作用エネルギーを計算し、
当該相互作用エネルギーについてリガンド座標データのコンフォメーションを微
調整しながら最適化する（ステップS-8）。リガンドのコンフォメーションの
10 微調整は、ステップS-7で算出されたリガンド座標を中心に、例えば、並進、
回転、または、シングルボンドまわりの角度をRMSDで0.3を越えない程度
に座標変化などをさせることで、行ってもよい。

ここで、リガンド座標データのコンフォメーションの微調整は例えばランダム
サーチで最適化することが好ましい。なお、ランダムサーチでは、以下の1)～
15 3)の項目に従って、タンパク質の活性部位とリガンドとの微小変化を例えば8
000回行い、最適エネルギー「U最適」が最小になるようにする。

1) 回転可能な結合のうち最大で5つを乱数で選び、結合ごとにランダムに±1
0.0°の範囲内で回転させ、リガンドのコンフォメーションを換える。この過
程を例えば3回に一度行う。

20 2) x、y、z軸方向それぞれにおいて、ランダムに±1.0Åの範囲内でリガ
ンドの並進運動を行う。この過程を例えば2回に一度行う。

3) 回転中心座標それぞれにおいて、ランダムに±1.0Åの範囲内で回転中心
座標を移動させ、さらに3次元方向の角度それぞれに対して、ランダムに±5.
0°の範囲内でリガンドの回転運動を行う。この過程を例えば5回に一度行う。

25 つぎに、本発明は、リガンド座標データのコンフォメーションを大きく変動し
て、ステップS-5から再スタートを行い、ステップS-8までを繰り返して最
適化を繰り返し行う（ステップS-9）。なお、コンフォメーション改変は、ス

ステップS-7で算出されたリガンド座標を中心に、例えば、並進、回転、または、シングルボンドまわりの角度をR SMDで0.3以上になるよう座標変化などをさせることで、行ってもよい。

ここで、リガンドのコンフォメーションを大きく動かした最適化は、例えばステップS-8で最適化したエネルギー「U最適」でのコンフォメーションに対し、
5 回転可能な結合をランダムに5つ選び、原子タイプごとに決められた回転角度間隔に従ってランダムに回転させる。その後ステップS-5以降の過程を、例えば5000回繰り返す。

また、リガンドのコンフォメーションを変化させた後、リガンドの内部エネルギー「U内部」を計算し、その値が500.0以上のときは、その後の計算をスキップし、次のリガンドコンフォメーションを発生させるようにしてもよい。
10

つぎに、本発明は、ステップS-4～ステップS-9までの過程を、ステップS-3で用意した複数の構造変化後タンパク質座標データに対して行い、最適なタンパク質とリガンドとの複合体座標、および最適エネルギー「U最適」を算出
15 する（ステップS-10）。

つぎに、本発明は、以上の過程を、ステップS-1で用意した化合物データベース中の全てのリガンドに対して行い、化合物データベース中から該当タンパク質と結合する可能性のあるリガンドを選択する（ステップS-11）。

以上、本発明の基本原理について説明したが、本発明は、タンパク質の誘導適合を反映するパラメータおよび／または構造変化した立体構造座標を分子動力学計算方法を用いて算出する場合、該当タンパク質の立体構造に対し基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、そのゆらぎの大きさを拘束条件として、分子動力学計算を行うことで、タンパク質の立体構造をエネルギー最適構造より大きく離れないようにして分子動力学計算を行ってもよい。
20

また、本発明において、本分子動力学計算方法による分子動力学計算は、例えば、基準振動計算より主鎖原子の2面角のゆらぎの値を算出し、当該ゆらぎ値を以下の数式2または数式3のように分子動力学計算における力の定数Kの部分に
25

入れてもよい。

$$E_{rot} = K_{rot} (\phi - \phi_0)^2 \quad \dots \text{(数式 2)}$$

- 5 E_{rot} はタンパク質の立体構造中において主鎖原子の 2 面角のエネルギーを示す。 ϕ は主鎖原子の 2 面角である。 ϕ_0 は主鎖原子の 2 面角の標準値である。ここで、 K_{rot} の値が大きい場合は、 ϕ は ϕ_0 に拘束される。

$$E_{pos} = K_{pos} (r - r_0)^2 \quad \dots \text{(数式 3)}$$

10

E_{pos} はタンパク質の立体構造中において主鎖原子の位置のエネルギーを示す。 r は主鎖原子の座標である。 r_0 は主鎖原子の座標の標準値である。ここで、 K_{pos} の値が大きい場合は、 r は r_0 に拘束される。

- また、本発明は、リガンドとタンパク質との相互作用を評価する際の目的関数
 15 (相互作用関数) として、従来の相互作用エネルギー関数に、タンパク質の動的性質を表現する動的性質関数を「弾性エネルギー」として加えてもよい。これにより、タンパク質の立体構造座標から相互作用エネルギーを高速に算出するとともに、タンパク質の動的挙動に関する物理化学的性質を明確に描写することができる。

- 20 ここで、「弾性エネルギー」として、タンパク質の局所的な柔らかさを考慮し、以下の数式 4 に示す関数「U 衝突」を適応してもよい。なお、以下の例においては、活性部位における動的挙動の少ない側鎖原子及び主鎖原子のみの i 番目の原子とリガンドの j 番目の原子との原子間距離 R が衝突距離「 R 衝突 (i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。

$$25 \quad U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

... (数式4)

Mは衝突を不許可とする活性部位の原子の数であり、Nはリガンドの原子の数である。また、「K衝突」は1000.0であることが好ましい。「R衝突(i, j)」は、活性部位のi番目の原子とリガンドのj番目の原子それぞれのV a n d e r W a a l s半径の和とする。

ここで、活性部位の各原子に対し、衝突を許す重み付けw(i)が定義された場合、以下の数式5に示す関数「U衝突」を用いる。ただし、w(i)は0～1の範囲の実数とする。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\phi(i, j) = w(i) * K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

... (数式5)

Mは活性部位の原子の数であり、Nはリガンドの原子の数である。また、「K衝突」は1000.0であることが好ましい。「R衝突(i, j)」は、活性部位のi番目の原子とリガンドのj番目の原子それぞれのV a n d e r W a a l s半径の和とする。

また、「弾性エネルギー」として、以下の数式6に示す関数を用いて定義することも可能である。

$$E_v = w(\text{hard shape region}), \\ E = 0(\text{soft shape region})$$

... (数式6)

ここで、「hard shape region」は、タンパク質の立体構造中、動的挙動の小さい部分のことを指し、「soft shape region」は、動的挙動の大きい部分のことを指す。また、Wは定数で100であることが好ましい。

また、本発明は、タンパク質の動的性質関数としてタンパク質の基準振動解析

結果または二次構造判定結果を用いてもよい。なお、二次構造判定においては、タンパク質のヘリックスやシート部分では揺らぎは小さいと考え、それ以外では揺らぎは大きいと考え、相互作用の評価関数、分子動力学計算の拘束条件に適應する。

5 また、本発明によれば、計算された該当タンパク質（タンパク質座標データ）が代表的な複数の立体構造座標である場合や、該当タンパク質の立体構造が例えば核磁気共鳴スペクトルの解析結果のような複数の立体構造座標である場合についても、該当タンパク質と結合するリガンドの探索について、複数座標すべてを全自動的にかつ短時間で同等に評価することができる。

10 また、本発明によれば、（１）タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択し、（２）選択された構造変化後タンパク質座標データから、リガンドと重ね合わせを行う空間点を指定し、（３）指定された空間点と、リガンドのリガンド座標データとを用いて、タンパク質とリガンド
15 とが結合した場合の相互作用関数を計算し、（４）計算された相互作用関数に基づいて当該タンパク質と結合するリガンドを評価する。これにより、リガンドのフレキシビリティだけでなく受容体側のフレキシビリティも考慮して、Induced-Fit型の受容体タンパク質に結合するリガンドを効率的に精度よく探索することができるという効果を奏する。

20 また、本発明によれば、タンパク質とリガンドとの結合に大変重要であるタンパク質の動的挙動を反映したパラメータを取得し、タンパク質の動的挙動を反映したリガンドとの新規な相互作用評価関数を用いて、該当タンパク質の立体構造と結合する新規リガンドを予測することができる。これにより、従来法と比較して、より信頼性の高いかつ医薬品設計等に適したタンパク質の立体構造を、世界
25 中で解析されている大量のゲノム配列に対しても対応するスピードで、構築することができる。従来、in silico スクリーニングにおいては、タンパク質とリガンドとの相互作用に重要な誘導結合を充分に取り扱うことのできるア

ルゴリズムが見出されていなかったが、本発明では、タンパク質とリガンドとの相互作用エネルギー関数に、基準振動計算結果、もしくは二次構造予測から得られるタンパク質の「ゆらぎ」を表すパラメータを簡易にとりこむ計算式を導入することで、誘導結合を充分に取り扱うことができる。

- 5 さらに、分子動力学シミュレーションにおいては、本発明では、該当タンパク質の動的挙動を反映したパラメータとリガンドとの相互作用評価関数に関して、該当タンパク質についての基準振動計算を行い、その計算結果を分子動力学計算に反映させることを特徴とする。従来、タンパク質の動的挙動のシミュレーションを行うためには分子動力学計算を用いていたが、従来法で該当タンパク質に対して分子動力学計算を行うと、タンパク質立体構造はX線やNMR等で解析された座標と大きくズレる。こうしたズレはタンパク質の動的挙動の物理化学的描写を含んでいるが、NMR等で示される動的挙動の実験的な結果と矛盾する挙動となる場合があり、必ずしも精度の高いシミュレーションとならないことが多い。そこで、分子動力学計算を行う際には、タンパク質の立体構造をある程度固定してシミュレーションを行う必要があり、本発明において、分子動力学計算におけるエネルギー関数中で主鎖原子の2面角に拘束をかける手法を開発した。さらに、2面角の拘束条件としては、そのパラメータとして予め該当タンパク質の基準振動計算を行い、主鎖原子の2面角のゆらぎを算出し、そのゆらぎの大きさにより例えばゆらぎの大きい部分は拘束条件を緩め、ゆらぎの小さい部分は拘束条件を強めるパラメータとして用いることとした。よって、本発明によれば、こうした条件でタンパク質の分子動力学シミュレーションを行うことで、精度よく動的挙動を描写することができる。加えて、こうして算出された分子シミュレーションからタンパク質の動的挙動を描写した座標を取得することができ、これを利用することで、さまざまなリガンド結合部位の形状を用いたリガンド探索を行うことができる。
- 10
15
20
25

これらの結果、本発明によれば、今までの *in silico* スクリーニングでは見出すことができなかった新規なリガンドを発見することを可能にすると

ともに、今までは長時間を必要とする分子シミュレーションでしか解析できなかった「誘導結合」を含めたタンパク質ーリガンドとの相互作用解析を短時間で行うことを可能にした。

本発明では、既存ソフトウェアよりも誘導結合現象をより深く考慮した “in
5 silico screening” に対応可能とし、誘導結合現象と疎水相
相互作用の正しい理解のもと単純化している。本発明は単純化されているので、自
動化により多くのターゲットタンパク質を処理可能とする。その結果、例えば1
00万以上の化合物データベースから、新規で、もっともらしい化合物を探索す
ることができるので、実験では対応できない規模のデータベースから、もっとも
10 らしい化合物を現実的な時間内に探索することができる。

また、本発明によれば、タンパク質ーリガンドとの相互作用解析を短時間で行
うことが可能になるので、例えば代謝、毒性の原因となる数多くのタンパク質と
薬物との相互作用解析が可能となり、in silicoでの薬物の代謝、毒性
予測を行うことができる。

15 本発明において、リガンドとして取り扱うことのできる分子は、使用するリガ
ンドの種類や数を限定しないため、蛋白質、ペプチド、DNA、薬剤成分、金属、
イオン、糖類、核酸成分、ホルモンを含む全ての物質を当該リガンドと見なすこ
とができる。本発明によって、具体的に農薬、医薬品等の分子設計を行うことが
できる。

20 また、リガンドとタンパク質との相互作用エネルギー評価関数には、従来、ド
ッキング法では静電エネルギー項、van der waals項、さらにはソ
フトドッキング法等に見られる動的挙動を表現するための調整項が主に用いられ
ているが、本発明においてはタンパク質とリガンドとの相互作用中にはソフトド
ッキング法等に見られる動的挙動を表現するための調整項を用いる代わりに古典
25 力学で用いられている弾性衝突の理論を適応し、タンパク質とリガンドとの相互
作用に関して、その物理化学的性質をより明確にした。このことにより、タンパ
ク質の構造変化と相互作用との関係を得ることができ、リガンドの機能の理解を

迅速かつ正確に行うための手助けとなる。

尚、本発明で利用するタンパク質の立体構造は、X線結晶構造解析等により3次元座標が決定されたタンパク質の立体構造以外に、タンパク質の経験的なモデリング法（特にホモロジーモデリング法やスレッディング法など）を利用して作成した立体構造座標も適応することができる。

[システム構成]

ここでは、本発明が適用される本システムの構成について、第116図を参照して詳細に説明する。第116図は、本発明が適用される本システムの構成の一例を示すブロック図であり、該構成のうち本発明に関係する部分のみを概念的に示している。

第116図に示すように、本システムは、概略的に、タンパク質と結合する物質（リガンド）を評価し、選定するリガンド探索装置100と、リガンド立体構造データやタンパク質立体構造データなどに関する外部データベースや各種の外部プログラムなどを提供する外部システム200とを、ネットワーク300を介して通信可能に接続して構成されている。

ネットワーク300は、リガンド探索装置100と外部システム200とを相互に接続する機能を有し、例えばインターネットやLANなどである。

外部システム200は、ネットワーク300を介して、リガンド探索装置100と相互に接続され、利用者に対してリガンド立体構造データやタンパク質立体構造データなどに関する外部データベースや各種の外部プログラムを実行するウェブサイトを提供する機能を有する。ここで、外部システム200は、WEBサーバやASPサーバ等として構成してもよく、そのハードウェア構成は、一般に市販されるワークステーション、パーソナルコンピュータ等の情報処理装置およびその付属装置により構成してもよい。また、外部システム200の各機能は、外部システム200のハードウェア構成中のCPU、ディスク装置、メモリ装置、入力装置、出力装置、通信制御装置等、およびそれらを制御するプログラム等により実現される。

リガンド探索装置 100 は、概略的に、リガンド探索装置 100 の全体を統括的に制御する CPU 等の制御部 102 と、通信回線等に接続されるルータ等の通信装置（図示せず）に接続される通信制御インターフェース部 104 と、各種のデータベースやファイルなどを格納する記憶部 106 と、入力装置 112 や出力装置 114 に接続される入出力制御インターフェース部 108 と、を備えて構成されており、これら各部は任意の通信路を介して通信可能に接続されている。さらに、リガンド探索装置 100 は、ルータ等の通信装置および専用線等の有線または無線の通信回線を介して、ネットワーク 300 に通信可能に接続されている。

記憶部 106 に格納される各種のデータベースやテーブルやファイル（リガンド座標データベース 106a ～ リガンド評価結果ファイル 106f）は、固定ディスク装置等のストレージ手段であり、各種処理に用いる各種のプログラムやテーブルやファイルやデータベースやウェブページ用ファイルなどを格納する。

これら記憶部 106 の各構成要素のうち、リガンド座標データベース 106a は、リガンド座標データを格納するリガンド座標データ格納手段である。タンパク質座標データベース 106b は、タンパク質座標データを格納するタンパク質座標データ格納手段である。構造変化後タンパク質座標データファイル 106c は、後述する構造変化後タンパク質座標データ選択部 102a により選択された構造変化後タンパク質座標データを格納する構造変化後タンパク質座標データ格納手段である。指定空間点ファイル 106d は、後述する空間点指定部 102b により指定された空間点に関する情報を格納する指定空間点格納手段である。相互作用関数計算結果ファイル 106e は、後述する相互作用関数計算部 102c により計算された相互作用関数の計算結果に関する情報を格納する相互作用関数計算結果格納手段である。リガンド評価結果ファイル 106f は、後述するリガンド評価部 102d により評価されたリガンドの評価結果に関する情報を格納するリガンド評価結果格納手段である。

通信制御インターフェース部 104 は、リガンド探索装置 100 とネットワーク 300（またはルータ等の通信装置）との間における通信制御を行う。すなわ

ち、通信制御インターフェース部 104 は、他の端末と通信回線を介してデータを通信する機能を有する。

入出力制御インターフェース部 108 は、入力装置 112 や出力装置 114 の制御を行う。ここで、出力装置 114 としては、モニタ（家庭用テレビを含む）
5 の他、スピーカ等を用いることができる（なお、以下においては出力装置 114 をモニタとして記載する場合がある。）。また、入力装置 112 としては、キーボードやマウス、マイクなどを用いることができる。また、モニタも、マウスと協働してポインティングデバイス機能を実現する。

制御部 102 は、OS (Operating System) 等の制御プログラム、および所要データを格納するための内部メモリを有し、これらのプログラム等により種種の処理を実行するための情報処理を行う。制御部 102 は、機能概念的に、構造変化後タンパク質座標データ選択部 102a と、空間点指定部 102b と、相互作用関数計算部 102c と、リガンド評価部 102d と、を含んで構成されている。

15 これら制御部 102 の各構成要素のうち、構造変化後タンパク質座標データ選択部 102a は、タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択する構造変化後タンパク質座標データ選択手段である。空間点指定部 102b は、構造変化後タンパク質座標データ選択部 102a により選
20 択された構造変化後タンパク質座標データから、リガンドと重ね合わせを行う空間点を指定する空間点指定手段である。

相互作用関数計算部 102c は、空間点指定部 102b により指定された空間点とリガンドのリガンド座標データとを用いて、タンパク質とリガンドとが結合した場合の相互作用関数を計算する相互作用関数計算手段である。ここで、相互作用関数計算部 102c は、第 117 図に示すように、相互作用関数最適化部 102c1 と、相互作用エネルギー最適化部 102c2 と、をさらに含んで構成されている。相互作用関数最適化部 102c1 は、相互作用関数のスコアが最大に

なるように最適化する相互作用関数最適化手段である。相互作用エネルギー最適化部102c2は、相互作用関数最適化部102c1により相互作用関数を最適化した後に、重ねあわせたリガンドに対してタンパク質との相互作用エネルギーを計算し、当該相互作用エネルギーについてリガンド立体構造データのコンフォメーションを微調整しながら最適化する相互作用エネルギー最適化手段である。

再び第116図に戻り、リガンド評価部102dは、相互作用関数計算部102cにより計算された相互作用関数に基づいて当該タンパク質と結合するリガンドを評価するリガンド評価手段である。ここで、リガンド評価部102dは、第118図に示すように、再評価部102d1をさらに含んで構成されている。再評価部102d1は、相互作用エネルギー最適化部102c2により最適化した後に、リガンド立体構造データのコンフォメーションを大きく変動させた後、再度、相互作用関数計算部102cを実行し、相互作用関数計算部102cにより計算された相互作用関数に基づいて当該タンパク質と結合するリガンドの再評価を行う再評価手段である。

15 なお、これら各部によって行われる処理の詳細については、後述する。

[システムの処理]

ここでは、上述のように構成された本実施の形態における本システムのメイン処理の一例について、第115図などを参照して詳細に説明する。第115図は、本実施形態における本システムのメイン処理の一例を示すフローチャートである。

20 ここでは、第115図を参照して、タンパク質立体構造と誘導適合を利用したリガンドの探索に関して説明する。

まず、リガンド探索装置100は、制御部102の処理により、3次元座標を含むリガンドのデータベースを用意し、用意したデータベースを記憶部106のリガンド座標データベース106aに格納する（ステップS0）。ここで、リガンドのデータベースとしては、例えば、ACD等のような市販化合物データベースや、化合物を描いて収集した仮想化合物データなどを用いてもよい。なお、リガンドのデータベースは、分子力学法等を用いて3次元化することが望ましい。

つぎに、リガンド探索装置 100 は、制御部 102 の処理により、ステップ S
0 で用意したリガンドデータベースから、特定リガンドを探索するための標的タンパク質の立体構造を選択し、当該選択したタンパク質の立体構造データ（3次元座標）を入手し、記憶部 106 のタンパク質座標データベース 106b に格納
5 する（ステップ S10）。なお、3次元座標には、公共データベースである PDB やホモロジーモデリング法等で作成した立体構造座標を用いることが望ましい。

つぎに、リガンド探索装置 100 は、構造変化後タンパク質座標データ選択部 102a の処理により、ステップ S10 により選択された標的タンパク質の基準振動計算を行い、主鎖原子の位置のゆらぎと 2 面角のゆらぎを求める（ステップ
10 S20）。具体的には、まず、ステップ S10 で定められた標的タンパク質の動的挙動を表すパラメータを基準振動解析法による計算結果のデータベースから取得する、もしくは当該パラメータを二次構造判定計算を行って取得する。

まず、ステップ S20 において、タンパク質の動的挙動を表すパラメータを基準振動解析法により取得する方法について説明する。基準振動解析法とは、ポテンシャルエネルギーを変位の二次関数として近似し、運動方程式を厳密に解き、
15 最適化構造の周りの微小な振動を解析する方法である。解くべき運動方程式は下記の式（1）または式（2）である。なお、基準振動解析法の詳細については、

「Wilson, E. B., Decius, J. C., and Cross, P. C. 1955. Molecular Vibration. McGraw-Hill.」に記載されている。
20

$$\left(\sum_j T_{ij} U_{jk} \right) \omega_k^2 = \sum_j V_{ij} U_{ik} \quad \dots (1)$$

$$TU\Lambda = VU$$

$$\Lambda_{ij} = \omega_i \delta_{ij}, U^T TU = (\delta_{ij}) \quad \dots (2)$$

ここで、 ω_k は固有値、 U_{ik} は固有ベクトルであり、 δ_{ij} はクロネッカーのデルタである。また、 T_{ij} と V_{ij} はそれぞれ、運動エネルギー E_k とポテンシャルエネルギー V に関係し、下記の式(3)および式(4)の通りである。

$$E_k = \frac{1}{2} \sum_{i,j} T_{ij} \dot{q}_i \dot{q}_j \quad \dots (3)$$

$$U_{vdw} = K_{vdw} \sum_{i,j(>i+2)} \left\{ \left(\frac{3.8}{D_{i,j}} \right)^{12} - \left(\frac{3.8}{D_{i,j}} \right)^6 \right\} \quad \dots (4)$$

ここで、 q_i は振動の自由度に対応した座標である。また、 q_i' (式(3)における「 q_i ドット」を意味する) は q_i の時間による微分である。また、 q_j は、下記の式(5)の通りである。

$$q_j = q_j^0 + \sum_k A_{jk} Q_k \quad \dots (5)$$

ここで、 A_{jk} は、集団運動 Q_k と個々の原子運動 q_j とを結ぶ係数である。また、 q_j^0 は最適化座標である。ただし、 Q_k は、下記の式に示す基準振動である。

$$Q_k = \alpha_k \cos(\omega_k t + \delta_k)$$

ここで、 α_k と δ_k は初期条件で定められる。

つぎに、ステップS20において、参照タンパク質に対して、上記で得られた固有値および固有ベクトルを用いて、ある温度・ある固有値での各C α 原子の位置ゆらぎを計算し、このゆらぎの値をC α が含まれるアミノ酸のゆらぎの値とする。また、目的タンパク質の各アミノ酸のゆらぎの値は、ステップS50におけるアライメントを利用して、目的配列と参照配列の比較から対応するアミノ酸残

基ペアにおいて、目的タンパク質のゆらぎの値として参照タンパク質と同一のものを当てはめておく。なお、ゆらぎの値を求められなかったものについては、予め設定した値をあてはめる。こうして得た目的タンパク質の各アミノ酸のゆらぎの値を目的タンパク質の動的な挙動を表すパラメータとする。

- 5 また、ステップS 2.0において、タンパク質の動的挙動を表すパラメータを二次構造判定計算を行って取得する方法について説明する。二次構造判定はタンパク質の立体構造座標から計算される。なお、ソフトウェアとしては、DSSP、STRIDE等が好ましいが、基本的にはタンパク質の主鎖のねじれ角と水素結合パターンから判別される方法を用いればよい。ここで、「DSSP (Dic
- 10 tionary of protein secondary structure of protein)」とは、PDB書式のファイルを入力ファイルとして、主鎖の水素結合パターンと、内部回転角等を解析し α ヘリックスと β シートとを判定するソフトウェアである。DSSPの詳細は、「Kabsch, W. &
- 15 Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers, 22:2577-2637」に記載されている。また、「STRIDE (Protein secondary structure assignment from a
- 20 tomic coordinate)」とは、PDB書式のファイルを入力ファイルとして、主鎖の水素結合パターンと、内部回転角等を解析し α ヘリックスと β シートとを判定するソフトウェアである。STRIDEの詳細は、「Fris
- 25 hman, D & Argos, P. (1995) Knowledge-based secondary structure assignment. Proteins: structure, function and genetics, 23, 566-579」に記載されている。

参照タンパク質に対して、上記ソフトウェア等を用いて二次構造計算を行い、

各アミノ酸がとる α ヘリックス構造、 β シート構造、ループ構造を判定する。目的タンパク質の各アミノ酸の二次構造は、ステップS50におけるアライメントを利用して、目的配列と参照配列の比較から対応するアミノ酸残基ペアにおいて、目的タンパク質の二次構造判定として参照タンパク質と同一のものを当てはめておく。二次構造判定を求められなかったものについては、予め設定しておいた結果をあてはめる。こうして得た目的タンパク質の各アミノ酸の二次構造判定結果を目的タンパク質の動的な挙動を表すパラメータとする。

ここで、ステップS20において、目的タンパク質の動的挙動を表すパラメータとしては、参照タンパク質の基準振動解析法により取得した計算結果を用いることが好ましい。なお、該当計算結果は別途データベースとして保存されているものを使用する。また、二次構造判定計算結果は、基準振動解析が行われていない参照タンパク質を用いる際に基準振動解析計算の代用として使用する。

再び第115図のメイン処理の説明に戻り、リガンド探索装置100は、構造変化後タンパク質座標データ選択部102aの処理により、ステップS20において求めた標的タンパク質のゆらぎを拘束条件として用いた分子動力学計算を行う（ステップS30）。

具体的には、まず、下記の式6に示す主鎖の位置拘束エネルギー「U位置」を導入し、初期の受容体骨格の変動を抑えながらAPRICOT[Yoneda S. & Umeyama H. (1992) Free energy perturbation calculations on multiple mutation bases J. Chem. Phys. 97, 6730-6736]を用いて最小化（条件：温度300K、受容体の表面から水分子が最低2分子配置できる箱状水槽、力場：AMBER [S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, & P. Weiner (1984) A new force field for molecular mechanical simulati

on of nucleic acids and proteins J.
Am. Chem. Soc. 106, 765-784]) を行う。

$$\text{U位置} = \text{K位置} * R^2$$

5 . . . (6)

ここで、「K位置」は例えば300.0とし、Rは基準座標からのずれとする。
つぎに、APRICOTに、下記の式7に示す二面角拘束エネルギー「U二面角」を導入して、最小化した受容体のMD計算（条件：温度300K、受容体の表面から水分子が最低2分子配置できる箱状水槽、力場：AMBER）を行う。

10

$$\text{U二面角} = \text{K二面角} * (\theta - \theta_{\text{平衡}})^2$$

. . . (7)

15

θ は二面角（単位rad）である。「K二面角」には、最大値と最小値を指定することで、その範囲内で主鎖二面角揺らぎに対応するように各二面角に対して不均一な拘束がかかるようにする。以後、主鎖二面角を拘束しながら行うMDを二面角拘束MDと呼ぶことにする。

20

つぎに、二面角拘束MD計算によりタンパク質構造座標を入手するには、受容体動的構造のクラスタリングを行う。あらかじめ指定した活性部位に対して、MDの途中経過100fsごとの受容体を重ね合わせた構造及び初期構造の活性部位を母集団とする。まず初めに、クラスタリングすることにより側鎖の動的情報が失われる可能性が高いことから、側鎖の二面角 χ において母集団の $\alpha\%$ が平均角度 $\pm 20.0^\circ$ の範囲で保存されている全側鎖二面角を収集する。ただし、主鎖の根元に近い方から χ が保存されていないと判定された場合はそれ以降の χ は保存されていないものとする。

25

つぎに、収集した保存側鎖二面角をすべて網羅している構造を母集団から抽出する。そして、抽出した構造の類似性を比較するために全原子rms (root mean square)が β Å以下の場合、同一構造と判断して一方を削除、

最終的に選ばれた構造をもとに受容体動的構造クラスターを作成する。なお、保存されていなかった二面角 χ を構成する原子では、変動する可能性が高いことから活性部位・リガンド結合計算において衝突しても良いことにする。ただし、 α 、 β は定数である。

- 5 再び第 1 1 5 図のメイン処理の説明に戻り、リガンド探索装置 1 0 0 は、空間点指定部 1 0 2 b の処理により、標的タンパク質のリガンド結合部位に、リガンドを配置するための空間点群を指定する（ステップ S 4 0）。具体的には、ステップ S 3 0 で作成した複数のタンパク質立体構造座標のうち、1 つをランダムに選択する。ここで、タンパク質座標中の空間点は、例えば以下の（1）または（
- 10 2）などの方法で指定させる。

（1）ダミー原子の発生による空間点の指定

- リガンドとタンパク質との相互作用における水素結合に着目し、タンパク質中の水素結合サイトを空間点として指定する。水素結合における重要事項は距離と角度である。つまり、角度を計算するためには、水素結合ドナー（以後、ドナー
- 15 ）に水素原子が必要になる。

そこで、本実施形態では、活性部位及びリガンドに水素原子が含まれていない場合、以下の規則によりダミー水素原子を発生させる。

- 1) sp^2 軌道原子を中心とする正三角形状にダミー原子を発生させる（第 2 図）。すなわち、第 2 図に示すように、 sp^2 軌道原子の窒素原子（A）を中心とする正三角形の空いている位置にダミー水素原子（B）を発生させる。
- 20 2) sp^3 軌道原子では、水素結合を形成する距離にある場合、水素原子を共有するように回転できると考え、水素結合相互作用を計算するときには、距離のみを考慮する。このため、 sp^3 軌道原子にはダミー原子を発生させない。

- また、金属及び水では、活性部位・リガンド結合の仲介役となり得るので、相互作用する位置に以下の規則によりダミー原子を発生させる。
- 25 1) 鉄のような金属には、正八面体状にダミー原子を発生させる（第 3 図）。すなわち、第 3 図に示すように、亜鉛（A）を中心とする正八面体の空いている位

置にダミー原子 (B) を発生させる。

2) 水には、正四面体状にダミー原子を発生させる。

ただし、活性部位と相互作用している方向にはダミー原子を発生させなくてもよい。

5 (2) 構造活性相関情報を利用した空間点の指定

また、本実施形態では、リガンドの構造活性相関 (SAR) 情報に着目し、以下の (A) ~ (D) 項目を入力情報にすることにより、空間点を指定する。

(A) SAR 情報から得られた活性部位の原子 (以後、「A 原子」)。PDB 形式に従う。

10 (B) 「A 原子」と相互作用するであろうリガンドの原子タイプ (以後、「B タイプ」)。SYBYL の MOL 2 形式に従う。

(C) 「A 原子」と「B タイプ」との相互作用の強さ (以後、「C 強さ」)。

(D) 「A 原子」と「B タイプ」との相互作用する距離 (以後、「D 距離」) (単位は Å)。

15 なお、本実施形態では、(A) ~ (D) の入力情報に基づいてタンパク質中の活性部位内におけるリガンドの初期座標を利用し、以下の 1) ~ 4) に示す規則により空間点を作成してもよい。

1) 「A 原子」がドナーまたは金属及び水の場合 (SAR 情報の活性部位側の指定が水素結合ドナー、金属原子の場合) には、「(1) ダミー原子の発生による空間点の指定」で発生させたダミー原子の方向に対して「A 原子」から「D 距離」の位置及びその周囲を初期座標に選ぶ (第 4 図、第 5 図)。

2) 「A 原子」が sp^3 軌道原子の場合 (SAR 情報の活性部位側の指定が sp^3 軌道原子の場合) には、「A 原子」から「D 距離」の周囲を初期座標に選ぶ (第 6 図)。

25 3) 「A 原子」が水素結合アクセプター (以後、アクセプター) の場合 (SAR 情報の活性部位側の指定が水素結合アクセプターの場合) には、「A 原子」の結合延長上「D 距離」の位置及びその周囲を初期座標に選ぶ (第 7 図)。

4) その他の場合 (SAR情報の活性部位側の指定がその他の原子の場合) には、「A原子」を中心とする半径が「D距離」の球表面上の点を初期座標に選ぶ (第8図)。

ここで、上記1)～4)とは異なり、リガンドの初期座標を直接指定してもよい。

再び第115図のメイン処理の説明に戻り、リガンド探索装置100は、相互作用関数計算部102cの処理により、ステップS0で定められた1つのリガンドに対し、リガンドの各原子をステップS30で定められた空間点群に重ね合わせる (ステップS50)。具体的には、距離行列を用いたアライメント作成アルゴリズム (DALI) [Holm, L., & Sander, C. (1993) Protein Structure Comparison by Alignment of Distance Matrices J. Mol. Biol. 233, 123-138] を低分子用に改良した以下の(1)～(4)に示す手法により初期座標とリガンドとを重ね合わせる。

(1) 1つの「Bタイプ」には、リガンドの原子タイプが複数対応することが多い。そこで、乱数を用いて「Bタイプ」とリガンドの原子タイプで同一視できるペアを作成する。ただし、ペアにおいてリガンドの原子タイプが重複しないようにする。

(2) 「Bタイプ」には、ステップS40により複数の初期座標が含まれているので、初期座標も乱数を用いて選択する。

(3) 選択された初期座標とリガンドそれぞれの距離行列を作成し、以下の相互作用関数である $Sscore(i, j)$ を計算する。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp \left\{ - \left(d_{ij}^s - d_{ij}^c \right)^2 \right\} - \beta \right] / \frac{\left(d_{ij}^s + d_{ij}^c \right)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases}$$

ここで、 d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。 α は、該当タンパク質中の空間点群と化合物が完全に重なりあった場合に $Sscore(i, j)$ を最大値とするための定数である。 β は重なりと定義できる限界値を与えるための定数である。

また、 α は1.5、 β は0.8とするのが好ましい。

(4) (1) ~ (3) を複数回 (例えば10000回) 繰り返し、 $Sscore(i, j)$ が最大になるペアを探し、そのペア情報をもとにリガンドを初期座標に重ね合わせる。

つぎに、リガンド探索装置100は、相互作用関数計算部102cの処理により、ステップS50での重ね合わせに対して、ステップS20およびステップS30で定められた計算結果によりタンパク質の動的挙動を表すパラメータを取得し、該当パラメータを用いてリガンドとタンパク質との相互作用エネルギーを、リガンドのコンフォメーションを微調整しながら計算する (ステップS60)。つまり、ステップS50で重ね合わせたリガンドに対して、タンパク質との相互作用エネルギーを、コンフォメーションを微調整しながら最適化して計算する。リガンドのコンフォメーションの微調整は、ステップS50で算出されたリガンド座標を中心に、例えば、並進、回転、または、シングルボンドまわりの角度を RSMDで0.3を越えない程度の座標変化などをさせることで、行ってもよい。

ここで、リガンド座標データのコンフォメーションの微調整は、例えばランダ

ムサーチで最適化することが好ましい。ランダムサーチでは、以下の1)～3)の項目に従って、タンパク質の活性部位とリガンドとの微小変化を例えば8000回行い、最適エネルギー「U最適」が最小になるようにする。

1) 回転可能な結合のうち最大5つ乱数で選び、結合ごとにランダムに±10.0°の範囲内で回転させリガンドのコンフォメーションを換える。この過程を例えば3回に一度行う。

2) x、y、z軸方向それぞれにおいて、ランダムに±1.0 Åの範囲内でリガンドの並進運動を行う。この過程を例えば2回に一度行う。

3) 回転中心座標それぞれにおいて、ランダムに±1.0 Åの範囲内で回転中心座標を移動させ、さらに3次元方向の角度それぞれに対して、ランダムに±5.0°の範囲内でリガンドの回転運動を行う。この過程を例えば5回に一度行う。

ここで、最適エネルギー「U最適」は以下の式で定義する。なお、当該式の右辺に示した各エネルギー関数については以下で順に説明する。

$$U_{\text{最適}} = U_{\text{SAR}} + U_{\text{水素}} + U_{\text{疎水}} + U_{\text{スタッキング}} + U_{\text{衝突}} + U_{\text{内部}}$$

ここで、原子のVan der Waals半径及び原子間相互作用距離は、AMBER 99 [J. Wang, P. Cieplak & P. A. Kollam (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. 21, 1049-1074]、およびMM3パラメータ [Ma B., Lii J.-H., Allinger N. L. (2000) Molecular polarizabilities and induced dipole moments in molecular mechanics J. Com

put. Chem. 21, 813-825] を参考にした。

(a) SAR 情報に関するエネルギー関数「 U_{SAR} 」

SAR 情報に従う指標としてエネルギー U_{SAR} を定義する。

$$U_{\text{SAR}} = \sum_{i=1}^N \phi(i)$$

$$\phi(i) = K_{\text{SAR}}(i) * (R_{\text{SAR}}(i) - R)^2 - \delta$$

N は SAR 情報の数であり、R は「A 原子」からリガンド側の相互作用原子までの距離であり、 $K_{\text{SAR}}(i)$ は i 番目の「C 強さ」であり、 $R_{\text{SAR}}(i)$ は i 番目の「D 距離」である。また、 δ は例えば 20.0 である。

(b) 水素結合に関するエネルギー関数「 $U_{\text{水素}}$ 」

リガンドの 1 つのドナー（アクセプター）に対して 1 つだけ水素結合を形成すると考え、最短にある活性部位側のアクセプター（ドナー）を選び、水素を介した結合角 θ （第 9 図参照。ただし、複数の水素原子が付加されているドナー原子の場合には、最小の水素結合角を θ と定義する。）を算出して、次の条件により分岐して、 $\phi(i)$ を計算するように定義する。なお、第 9 図において、A はドナー、B は水素、C はアクセプター、 θ は水素結合角を示す。

$$U_{\text{水素}} = \sum_{i=1}^N \phi(i)$$

(1) ドナー原子が sp^3 軌道原子、または水素結合角 θ が $\pm 30.0^\circ$ 以内のとき

$$\text{If } R > R_{\text{水素}}, \quad \phi(i) = - \frac{K_{\text{水素}}(i)}{(R - R_{\text{水素}}(i) + 1.0)}$$

$$\text{Else,} \quad \phi(i) = - \frac{K_{\text{水素}}(i)}{(R_{\text{水素}}(i) - R + 1.0)}$$

(2) 水素結合角 θ が $\pm 30.0^\circ$ 以上のとき

$$\begin{aligned} \text{If } R > R_{\text{水素}}, \quad \phi(i) &= - \frac{K_{\text{水素}}(i)}{(R - R_{\text{水素}}(i) + 1.0) * \theta} \\ \text{Else,} \quad \phi(i) &= - \frac{K_{\text{水素}}(i)}{(R_{\text{水素}}(i) - R + 1.0) * \theta} \end{aligned}$$

5

Nはリガンドのドナー+アクセプターの数であり、Rは水素結合を形成する二原子間距離であり、「K水素(i)」及び「R水素(i)」は原子タイプごとに決めた水素結合の相互作用の強さ及び距離である。

(c) 疎水相互作用エネルギー関数「U疎水」

- 10 活性部位 (ALA、CYS、PHE、ILE、LEU、MET、PRO、VAL、TRP、TYRの側鎖。ただし、TYRの水酸基は除く) 及びリガンド (炭素原子) の疎水相互作用し得る原子に通し番号を付け、活性部位の i 番目とリガンドの j 番目との原子間距離 R がカットオフ以内にあるとき $\phi(i, j)$ を計算するように定義する。

$$15 \quad U_{\text{疎水}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\begin{aligned} \text{If } R > R_{\text{疎水}}(i, j), \quad \phi(i, j) &= - \frac{K_{\text{疎水}}(i, j)}{(R - R_{\text{疎水}}(i, j) + 1.0)} \\ \text{Else,} \quad \phi(i, j) &= -K_{\text{疎水}}(i, j) \end{aligned}$$

- 20 Mは活性部位の疎水相互作用し得る原子の数であり、Nはリガンドの疎水相互作用し得る原子の数であり、「K疎水(i, j)」及び「R疎水(i, j)」は原子タイプごとに決めた疎水相互作用の強さ及び距離である。また、カットオフは例えば 8.0 Å とする。

(d) スタッキングエネルギー関数「Uスタッキング」

- 25 活性部位及びリガンドの芳香環を形成する原子に通し番号を付け、活性部位においては芳香環の中心座標を算出した。活性部位の i 番目とリガンドの j 番目との原子間距離 R がカットオフ以内にあるとき、i 番目の原子が形成する芳香環の

- 中心座標を i' 、 j 番目の最短距離あり共に同じ芳香環を形成するリガンドの原子を j' としたとき、 $\angle i i' j = \theta_{i' j}$ 、 $\angle i' i j = \theta_{i j}$ 、 $\angle i i' j' = \theta_{i' j'}$ 、 $\angle i' i j' = \theta_{i j'}$ を算出し (第10図)、 $\theta_{i' j}$ と $\theta_{i j}$ とが $90.0^\circ \pm 10.0^\circ$ のとき、「R境界」と「 θ 境界」を求め次の条件により分岐して $\phi(i, j)$ を計算するように定義する。第10図において、 i' は活性部位の芳香環中心を示し、 i は活性部位の芳香環原子を示し、 j および j' はリガンドの芳香環原子を示す。

$$U_{\text{スタッキング}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

- 10 If R境界 < 0.0 ,

$$\phi(i, j) = -K_{\text{スタッキング}}(i, j) * R_{\text{境界}}$$

Else,

$$\phi(i, j) = -K_{\text{スタッキング}}(i, j) * \theta_{\text{境界}}$$

- 15 $R_{\text{境界}} = 1.0 - (R_{\text{スタッキング}}(i, j) - R)^2$

$$\theta_{\text{境界}} = |1.0 - \Theta|$$

$$\Theta = \frac{\pi}{180.0} * (\theta - 90.0)^2$$

- 20 Mは活性部位の芳香環を形成する原子の数であり、Nはリガンドの芳香環を形成する原子の数であり、「 $K_{\text{スタッキング}}(i, j)$ 」及び「 $R_{\text{スタッキング}}(i, j)$ 」は原子タイプごとに決めたスタッキングの強さ及び距離である。 π は円周率であり、 θ は $\theta_{i' j}$ と $\theta_{i j}$ において Θ が最小になる角度である。また、カットオフは例えば5.0 Åとする。

- 25 (e) 分子間衝突エネルギー関数 (弾性衝突エネルギー関数) 「U衝突」

「弾性エネルギー」として、タンパク質の局所的な柔らかさを考慮し、以下の式に示す関数「U衝突」を適応してもよい。活性部位における動的挙動の少ない

(保存されている) 側鎖原子及び主鎖原子のみの i 番目の原子とリガンドの j 番目の原子との原子間距離 R が衝突距離「 R 衝突 (i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

5

$$\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

M は衝突を不許可とする活性部位の原子の数であり、 N はリガンドの原子の数である。また、「 $K_{\text{衝突}}$ 」は 1000.0 であることが好ましい。「 $R_{\text{衝突}}(i, j)$ 」は活性部位の i 番目の原子とリガンドの j 番目の原子それぞれの Van der Waals 半径の和とする。

10

ここで、活性部位の各原子に対し、衝突を許す重み付け $w(i)$ が定義された場合、以下の式に示す関数「 $U_{\text{衝突}}$ 」を用いる。ただし、 $w(i)$ は 0~1 の範囲の実数とする。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

15

$$\phi(i, j) = w(i) * K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

20

M は活性部位の原子の数であり、 N はリガンドの原子の数である。また、「 $K_{\text{衝突}}$ 」は 1000.0 であることが好ましい。「 $R_{\text{衝突}}(i, j)$ 」は活性部位の i 番目の原子とリガンドの j 番目の原子それぞれの Van der Waals 半径の和とする。

(f) リガンド内部エネルギー「 $U_{\text{内部}}$ 」

25

回転可能な結合を微小に変動させていくと、誤差で結合が切れる恐れがあるために「 $\phi_{\text{結合長}}(i)$ 」を、また、リガンド内部で原子衝突が起こることも避けるために「 $\phi_{\text{衝突}}(i, j)$ 」を計算するように定義する。

$$U_{\text{内部}} = \sum_{i=1}^L \phi_{\text{結合長}}(i,j) + \sum_{i=1}^M \sum_{j=1}^N \phi_{\text{衝突}}(i,j)$$

$$\phi_{\text{結合長}}(i) = K_{\text{結合長}} * \{1000.0 * (R_{\text{結合長}}(i) - R_1)\}^2$$

$$\phi_{\text{衝突}}(i, j) = K_{\text{衝突}} * (R_{\text{衝突}} - R_2)^2$$

5

Lは回転可能な結合の数であり、Mはリガンドの原子数であり、Nはi番目の原子の非結合原子数である。また、「K結合長」は100.0であることが好ましい。「R結合長(i)」は初期構造の結合長である。「K衝突」は150.0であり、「R衝突」は2.2 Åであることが好ましい。また、R₁とR₂は二原子間距離とする。

10

再び第115図のメイン処理の説明に戻り、リガンド探索装置100は、相互作用関数計算部102cの処理により、ステップS50で定められたリガンドに対して、リガンド座標データのコンフォメーションを大きく変動して、ステップS50から再スタートを行い、ステップS60までを繰り返して最適化を繰り返して行う(ステップS70)。なお、コンフォメーション改変は、ステップS50で算出されたリガンド座標を中心に、例えば、並進、回転、または、シングルボンドまわりの角度をRMSDで0.3以上になるよう座標変化などさせることで、行ってもよい。

15

ここで、リガンドのコンフォメーションを大きく動かした最適化は、例えばステップS50で最適化したエネルギー「U最適」でのコンフォメーションに対して、回転可能な結合をランダムに5つ選び、原子タイプごとに決められた回転角度間隔に従ってランダムに回転させる。その後ステップS50、ステップS60の過程を、例えば5000回繰り返し行う。

20

また、リガンドのコンフォメーションを変化させた後、リガンドの内部エネルギー「U内部」を計算しその値が例えば500.0以上のときは、その後の計算をスキップし、次のリガンドコンフォメーションを発生させてもよい。

25

つぎに、リガンド探索装置100は、相互作用関数計算部102cの処理によ

り、ステップS 70まで得られた標的タンパク質とリガンドとの相互作用エネルギーを決定する（ステップS 80）。具体的には、ステップS 40からステップS 70まで「U最適」が最適値となる最適なタンパク質とリガンドとの複合体座標、および最適エネルギー「U最適」を算出する。

5 つぎに、リガンド探索装置100は、制御部102の処理により、ステップS 40に戻り、ステップS 0中の別のリガンドを選択し、各処理部の処理により、ステップS 80まで計算する（ステップS 90）。なお、ステップS 40からステップS 90まではステップS 0中の化合物データベース中のリガンド全てについて行う。

10 つぎに、リガンド探索装置100は、リガンド評価部102dの処理により、ステップS 0中でのリガンドに対し、ステップS 90において定められた相互作用エネルギーを比較し、標的タンパク質に結合すると予想されるリガンドを選択する（ステップS 100）。具体的には、ステップS 90まで評価されたタンパク質とリガンドとの複合体座標および最適エネルギー「U最適」に基づいて、ステップS 0中のデータベース中から該当タンパク質と結合する可能性のある化合物（リガンド）を選択する。

以上、本システムのメイン処理の説明を終了する。

以上、説明したように、リガンド探索装置100によれば、相互作用関数によって、該当タンパク質と結合する物質（具体的にはリガンド）を評価し、選定することができる。具体的には、該当タンパク質の立体構造に際し、基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、そのゆらぎの大きさを拘束条件として分子動力学計算を行うことで、タンパク質の立体構造をエネルギー最適構造よりおおきく離れないようにして分子動力学計算を行うことができる。また、タンパク質の動的挙動に関する物理化学的性質を明確に描写することができる。また、タンパク質の動的性質を表現する関数として基準振動解析結果またはタンパク質の二次構造判定結果を用いることができる。また、核磁気共鳴スペクトルの解析結果のような複数の立体構造座標である場合についても、該当タンパク質と

結合するリガンドの探索について、複数座標すべてを全自動的にかつ短時間で同等に評価することができる。

また、リガンド探索装置 100 によれば、任意の単数を含む複数鎖のタンパク質立体構造が与えられた場合において、該当タンパク質の立体構造から誘導適合を反映したパラメータおよび構造変化した立体構造座標を例えば基準振動計算方法や分子動力学計算方法よりあらかじめ算出し、当該パラメータおよび構造変化した立体構造座標を用いて該当タンパク質と別の物質が結合した場合の相互作用関数を定義し、当該相互作用関数によって該当タンパク質と結合する物質をコンピュータプログラムにより評価し、選定する。

また、リガンド探索装置 100 によれば、該当タンパク質に結合するリガンドを選択する際に (0) ~ (8) に示した一連の処理を全自動または手動的に行う。

(0) 化合物データベースからリガンドを 1 つ選択する。該当タンパク立体構造として、誘導適合を反映するパラメータを用いて動的挙動を考慮した複数の構造変化座標を用意し、ランダムに 1 つの構造を選択する。

(1) 重ね合わせを行う該当タンパク中の空間点を指定する。

(2) (0) で選択したリガンド中の原子と (1) で指定した空間点とのペアを重複がないようにランダムに選択する。

(3) 以下のスコア $Sscore(i, j)$ を計算する。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp \left\{ - (d_{ij}^s - d_{ij}^c)^2 \right\} - \beta \right] / \frac{(d_{ij}^s + d_{ij}^c)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases}$$

d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。 α は、該当タンパク質中の空間点群

と化合物が完全に重なりあった場合に $S\ score(i, j)$ を最大値とするための定数である。 β は重なりと定義できる限界値を与えるための定数である。

(4) (3) のスコアが最大になるように調整する。

5 (5) (4) で重ねあわせたリガンドに対してタンパク質との相互作用エネルギーをコンフォメーションを微調整しながら最適化計算する。

(6) リガンドのコンフォメーションを大きく動かして、(2) から再スタートを行い、(5) までを繰り返して最適化を行う。

10 (7) (1) ~ (6) までの過程を (0) で用意した複数の構造変化座標に対して行い、最適なタンパク質とリガンドとの複合体座標、最適エネルギー「 $U_{最適}$ 」を算出する。

(8) (1) ~ (7) までの過程を (0) で用意した化合物データベース中の全てのリガンドに対して行い、化合物データベース中から該当タンパク質と結合する可能性のあるリガンドを選択する。

15 また、リガンド探索装置 100 によれば、タンパク質の誘導適合を反映するパラメータおよび構造変化した立体構造座標を分子動力学計算方法を用いて算出する場合、該当タンパク質の立体構造に際し、基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、そのゆらぎの大きさを拘束条件として、分子動力学計算を行うことで、タンパク質の立体構造をエネルギー最適構造よりおおきく離れないようにして分子動力学計算を行う。

20 また、リガンド探索装置 100 によれば、リガンドとタンパク質との相互作用を評価する際の目的関数として、従来の相互作用エネルギー関数に、タンパク質の動的性質を表現する関数を弾性エネルギーとして加え、タンパク質の立体構造座標から相互作用エネルギーを高速に算出するとともに、タンパク質の動的挙動に関する物理化学的性質を明確に描写する。

25 また、リガンド探索装置 100 によれば、タンパク質の動的性質を表現する関数として基準振動解析結果またはタンパク質の二次構造判定結果を用いる。

また、リガンド探索装置 100 によれば、計算された該当タンパク質が代表的

な複数の立体構造座標である場合や、該当タンパク質の立体構造が例えば核磁気共鳴スペクトルの解析結果のような複数の立体構造座標である場合についても、該当タンパク質と結合するリガンドの探索について、複数座標すべてを全自動的にかつ短時間で同等に評価することを可能とする。

5 [実施例 1]

(二面角拘束MDおよびクラスタリングにおけるパラメータ定数の決定)

10 上述した実施形態におけるリガンド探索装置 100 を用いて、基準振動解析により二面角のゆらぎ値を計算した。なお、本実施例 1 においては、二面角のゆらぎ値を、分子動力学計算における拘束条件として、以下の式中の「K二面角」に適応した。

$$U_{\text{二面角}} = K_{\text{二面角}} * (\theta - \theta_{\text{平衡}})^2$$

15 θ は二面角 (単位 rad) である。実際には、「K二面角」の最大値と最小値を指定することにより、その範囲内で主鎖二面角揺らぎに対応するように各二面角に対して不均一な拘束がかかるようにした。そのため、本実施例 1 では、「K二面角」の適切な最大値と最小値を決定することを目的とした。

20 また、分子動力学計算後、構造変化した座標をクラスター解析し、代表構造を選択した。その際、あらかじめ指定した活性部位に対して、MD の途中経過 1000 frames ほどの受容体を重ね合わせた構造の活性部位及び初期構造の活性部位を母集団とした。具体的には、まず、クラスタリングすることにより側鎖の動的情報が失われる可能性が高いことから、側鎖の二面角 χ において母集団の $\alpha\%$ が平均角度 $\pm 20.0^\circ$ の範囲で保存されている全側鎖二面角を収集した。ただし、主鎖の根元に近い方から χ が保存されていないと判定された場合はそれ以降の χ は保存されていないものとした。つぎに、収集した全側鎖二面角 (保存側鎖二面角) をすべて網羅している構造を母集団から抽出した。つぎに、抽出した構造の類似性を比較するために、全原子 rms (root mean square

）が β Å 以下の場合には同一構造と判断して一方を削除した。そして、最終的に選ばれた構造をもとに受容体動的構造クラスターを作成した。ただし、保存されていなかった二面角 χ を構成する原子では、変動する可能性が高いことから、活性部位・リガンド結合計算において衝突しても良いことにした。なお、 α 、 β は定数であり、本実施例 1 では、適切な α 、 β を決定することも目的とした。

さらに、本実施例 1 では、リガンドと接触している活性部位において最も良い主鎖の動的構造を得ることも目的とした。そのため、rms (root mean square) を計算するときは、活性部位における主鎖原子 (N、C α 、C、O) の 4 原子のみを対象とした。

「K二面角」の最大値と最小値、およびクラスタリング定数の α と β は、NMR で解析された構造を再現できる値が適切であると考えた。そこで、まず、NMR で解析された構造のジヒドロ葉酸還元酵素 (DHFR、PDB code: 1LUD) の MODEL 1 を初期構造に基準振動解析を行い、ゆらぎ値を求め、その後、分子動力学計算を行った。そして、分子動力学計算後は、受容体動的構造クラスタリングまで行った。なお、1LUD (MODEL 1) に含まれていたリガンドの各原子から半径 6 Å 以内に含まれる受容体残基を活性部位と定義した。また、MD は 0 ~ 0.1 nsec までの結果を使用した。ここで、MD での拘束の最小値と最大値は 0 から 1000 まで (100 ごと) の数値、クラスタリングにおいて定数 α は 0 % から 90 % まで (10 % ごと) の数値、定数 β は NMR 構造平均値 rms を参考に 0.1 Å から 0.6 Å まで (0.1 Å ごと) の数値に対して網羅的に行って、1LUD における NMR 構造すべてと比較することにより、定数を決定した。

受容体動的構造クラスタリングにおける定数 β を決定するための参考として NMR 構造平均値を求めた。PDB (the Protein Data Bank) の NMR 構造のうち、受容体が単純タンパク質であり且つ 1 つの PDB ファイル内に記載されていた NMR 構造が 10 パターン以上あった。そのため、リガンドを含む 117 種類を対象に活性部位の NMR 構造平均値 rms を求めること

にした。

まず、MODEL 1において、リガンドの各原子から半径6 Å以内に含まれる受容体残基を活性部位と定義した。そして、MODEL 1以外の構造において、MODEL 1の活性部位とのrmsをそれぞれ求め、さらにその平均rmsを求めた。ここで、平均rmsが1.0 Å以上の場合は、明らかな動的構造と見なせるので、そのようなPDBファイルを対象からはずした。これにより、対象となるPDBファイルは71種類となった。そして、71種類の平均rmsをさらに平均化した値をNMR構造平均値rmsとした。このようにして得られたNMR構造平均値rmsは0.62となった。

10 「K二面角」の適切な最大値と最小値の決定、およびクラスタリングにおける定数 α と β の決定に関しては、各パラメータ値とNMR構造との比較を行った。

1 LUDには24種類のMODELが含まれており、また本実施例1ではMODEL 1を対象にしたので、これを除く23種類のMODELの活性部位を正解構造とした。計算の結果として出力された各受容体動的構造クラスターにおいて、各正解構造とrmsを計算した。なお、計算したrmsの中で最小のrmsを「RMS最小」とし、各受容体動的構造クラスターから得られた「RMS最小」の平均値をスコアとした。そして、このスコアが最小となるパラメータを採用することにした。

第11図に1 LUDのMODEL 1における基準振動解析結果を示し、第12図、第13図、第15図～第18図にスコアとパラメータとの比較の結果を示した。第11図では、二面角 ϕ （橙色A）、 ϕ （緑色B）の揺らぎの大きさを示した。揺らぎが0.0に近いほど分子動力学法（MD）計算において二面角拘束が強くなる。また、STRIDEによる二次構造判定で α -ヘリックス（赤色D）、 β -シート（青色D）を表示した。紫色Cは活性部位である。第12図において、基準振動解析の結果、定数 α は70%が良いという結果になった。しかし、一般性を持たせときに70%ではクラスタリングの精度が低下した場合もあったので、本実施例1では80%を定数 α の値にした。第13図、第15図～第18図にお

いて、クラスタリング定数を $\alpha = 80.0\%$ 、 $\beta = 0.4 \text{ \AA}$ に固定した。なお、黒色に近いほどスコアが小さいことを表わす。

- これらの結果より、スコアが小さくなる拘束条件としては第14図の値が最適であると判断された。これらの値の妥当性は、例えば、主鎖原子のみではなく、
- 5 C α 原子、側鎖原子、全原子について調査しても第14図のパラメータ値が最適であることが分かった。

[実施例2]

(拘束パラメータ有無による分子動力学計算の相違)

- 上述したリガンド探索装置100により算出された拘束パラメータを適応した
- 10 分子動力学計算を2.0 nsecまで行った。そして、活性部位の主鎖原子の動的挙動において、拘束パラメータを適応しない場合と比較して、構造がどの程度変化するかを調べた。

Case 1)

- ジヒドロ葉酸還元酵素 (1LUDのMODEL 1) を対象に検証した。検証結果を第19図～第21図に示した。なお、基準振動計算結果は上述した実施例1
- 15 で求めた値を適応した。

- 第19図では、1LUDのPDBファイル内に記載されている24種類の各モデル構造をMODEL 1と活性部位の主鎖原子においてrmsを計算し、その平均rmsを点線で表示した。二面角拘束があるとき (A) と、二面角拘束がない
- 20 とき (B) において、活性部位の主鎖原子の初期構造からのずれをrmsで表示した。

第20図に、二面角拘束なしでMDを計算させた時のNMR構造との比較結果を示した。第20図において、白色はNMR構造 (1lud) であり、黒色はMD構造 (1lud) である。

- 25 表1に、二面角拘束なしでMDを計算させた時のNMR構造との比較結果を示した。

57

(表 1)

	活性部位	全体
C α のみ	3.8903	0.2919
主鎖	3.8642	0.3335
全体	4.447	0.1398 RMS

第 2 1 図に、二面角拘束ありでMDを計算させた時のNMR構造との比較結果
5 を示した。

表 2 に、二面角拘束ありでMDを計算させた時のNMR構造との比較結果を示
した。

(表 2)

	活性部位	全体
C α のみ	0.6398	0.1194
主鎖	0.6933	0.1053
全体	1.2379	0.2157 RMS

Case 2)

ここでは、FAMS [Ogata K., Umeyama H. (2000) An automatic homology modeling method consisting of database searches
15 and simulated annealing J. Mol. Graphics Mod. 18, 258-272] によりモデリングした構造 (モデル構造) と X 線構造を初期構造に選び、初期構造及び拘束の有無に依存することを検証した。なお、リガンドの各原子から半径 10 Å 以内に含まれる受容体残基を活性部位と定義した。

20 cellular retinoic acid binding protein type II (CRABP-II) (PDB code: 1CBQ) の X 線構造 (立体構造) を利用した。また、参照タンパク質にホモロジー 32.1% の intestinal fatty acid binding protein (PDB code: 1ICM) を選び、第 2 2 図のアライメントでモデル構造を作成した。第 2 3 図、第 2 4 図、第 2 5 図に、X 線とモデルとの構造
25 比較結果を示した。

第 2 3 図には、1CBQ の立体構造 (X 線構造 (赤色 A) およびモデル構造 (

青色B))を示した。第24図には、第23図の緑色Cで示される物質である6
- (2, 3, 4, 5, 6, 7-hexahydro-2, 4, 4-trimeth
hyl-1-methyleneinden-2-yl)-3-methylh
exa-2, 4-dienoic acidの構造を示した。第25図には、1
5 CBQのX線構造とモデル構造の相違をrmsで表示した。

第26図は、1CBQのX線構造の基準振動解析の結果を示す図であり、第2
7図は1CBQのモデル構造の基準振動解析の結果を示す図である。第26図お
よび第27図では、二面角 ϕ (橙色A)、 ϕ (緑色B)の揺らぎの大きさを示し
た。揺らぎが0.0に近いほど分子動力学法(MD)計算において二面角拘束が
10 強くなる。また、STRIDEによる二次構造判定で α -ヘリックス(赤色D)、
 β -シート(青色D)を表示した。紫色Cは活性部位である。

第28図には、1CBQのX線構造とモデル構造の分子動力学法(MD)計算
の結果を示した。X線構造の活性部位の主鎖原子とのrmsを求めた。第28図
において、Aは初期構造がX線構造で二面角拘束なし、Bは初期構造がX線構造
15 で二面角拘束あり、Cは初期構造がモデル構造で二面角拘束なし、Dは初期構造
がモデル構造で二面角拘束あり、である。

Case 3)

FlavodoxinのX線構造(PDB code: 1J9G)を利用した。
また、参照タンパク質にホモロジー29.2%のflavodoxin(PDB
20 code: 1AHN)を選び、第29図のアライメントでモデル構造を作成し
た。第29図には、1J9Gおよび1AHNのアライメントを示した。

第30図には、1J9Gの立体構造(X線構造(赤色A)およびモデル構造(
青色B))を示した。第31図には、第30図における緑色Cで示される物質で
あるflavin mononucleotideの構造を示した。第32図に
25 は、1J9GのX線構造とモデル構造との相違をrmsで表示した。

第33図には1J9GのX線構造の基準振動解析の結果を、第34図には1J
9Gのモデル構造の基準振動解析の結果を示した。第33図および第34図にお

いて、二面角 ϕ (橙色A)、 ϕ (緑色B) の揺らぎの大きさを示した。揺らぎが 0. 0 に近いほど分子動力学法 (MD) 計算において二面角拘束が強くなる。また、STRIDE [8] による二次構造判定で α -ヘリックス (赤色D)、 β -シート (青色D) を表示した。紫色Cは活性部位である。

- 5 第35図には、1J9GのX線構造とモデル構造の分子動力学法 (MD) 計算の結果を示した。X線構造の活性部位の主鎖原子との rms を求めた。第35図において、Aは初期構造がX線構造で二面角拘束なし、Bは初期構造がX線構造で二面角拘束あり、Cは初期構造がモデル構造で二面角拘束なし、Dは初期構造がモデル構造で二面角拘束あり、である。

10 Case 4)

Matrix metalloproteinase-8 (MMP-8) のX線構造 (PDB code: 1MMB) を利用した。また、参照タンパク質にホモロジー55. 0%のMMP-3 (PDB code: 1B3D) を選び、第36図のアライメントでモデル構造を作成した。第36図には、1MMBおよび1B3D_Aのアライメントを示した。

- 15 第37図には、1MMBの立体構造 (X線構造 (赤色A) およびモデル構造 (青色B)) を示した。第38図には、第37図における緑色Cで示される物質である batimastat の構造を示した。第39図には、1MMBのX線構造とモデル構造との相違を rms で表示した。

- 20 第40図には1MMBのX線構造の基準振動解析の結果を、第41図には1MMBのモデル構造の基準振動解析の結果を示した。第40図および第41図において、二面角 ϕ (橙色A)、 ϕ (緑色B) の揺らぎの大きさを示した。揺らぎが 0. 0 に近いほど分子動力学法 (MD) 計算において二面角拘束が強くなる。また、STRIDE [8] による二次構造判定で α -ヘリックス (赤色D)、 β -シート (青色D) を表示した。紫色Cは活性部位である。

- 25 第42図には、1MMBのX線構造とモデル構造の分子動力学法 (MD) 計算の結果を示した。X線構造の活性部位の主鎖原子との rms を求めた。第42図

において、Aは初期構造がX線構造で二面角拘束なし、Bは初期構造がX線構造で二面角拘束あり、Cは初期構造がモデル構造で二面角拘束なし、Dは初期構造がモデル構造で二面角拘束あり、である。

5 Case 1) ~ Case 4) に示したとおり、拘束パラメータを適応した分子動力学計算結果は、拘束パラメータを適応しない場合と比較して、大きな構造変化は少ない。このことは古典力学を適応しているため大きな構造変化をしてしまう分子動力学法において、拘束パラメータを適応することで大きな構造変化を合理的に拘束することができ、理想的な構造座標を得ることが可能であるということを示している。また、ホモロジーが高ければ、FMASの構造構築精度も上がる。すなわち、X線に近い構造を得られるので、アミノ酸の数個異なるミューテ

10 ーションタンパク質にも本発明は利用できる。

[実施例3]

(タンパク質／リガンド複合体モデルの検証)

上述した実施形態におけるリガンド探索装置100により、該当タンパク質に結合するリガンドの複合体立体構造を予測した。本実施例3では、予測された複合体立体構造座標の予測精度を検証した。なお、当該検証には、複合体の立体構造が既知で、リガンドの有無もしくはリガンドの種類により活性部位の形が異なるInduced-Fit型のタンパク質を用いた。また、リガンドの各原子から半径10 Å以内の残基をタンパク質の活性部位と定義した。また、X線構造またはNMR構造を初期構造に選んだMDでは、ほぼ一定の構造を保ち続けることが分かったのでMDを1.0 nsecまで行うことにした。ただし、水素原子を除いて計算した。複合体モデル構築は、上述した実施形態に従って行った。

15 20

Case 1)

ジヒドロ葉酸還元酵素 (DHFR) である1BZFと1LUDとはホモロジーが100.0%でかつ結合しているリガンドが異なることにより活性部位の形が異なる。そこで、1BZF (MODEL 18) を初期構造として選択し、リガンドとして2, 4-diamino-5-(3, 4, 5-trimethoxy-

25

benzyl)-pyrimidin-1-ium (第49図)を用い、上述した実施形態におけるリガンド探索装置100によってタンパク質/リガンド複合体モデルを作成し、正解構造である1LUD (MODEL 4)と比較することで検証した (第43図)。第43図には、ジヒドロ葉酸還元酵素の立体構造を示した。第43図では、1LUD (MODEL 4) 受容体 (緑色A) とリガンド (赤色B)、1BZF (MODEL 18) の受容体 (青色C) とリガンド (水色D) を示した。

第44図には、1BZFの基準振動計算解析を示した。第44図では、二面角 ϕ (橙色A)、 ϕ (緑色B) の揺らぎの大きさを示した。揺らぎが0.0に近いほど分子動力学法 (MD) 計算において二面角拘束が強くなる。また、STRIDEによる二次構造判定で α -ヘリックス (赤色D)、 β -シート (青色D) を表示した。紫色Cは活性部位である。

第45図、第47図に1BZFを用いた拘束二面角分子動力学の結果を示した。第45図には、正解構造1LUD (MODEL 4) の活性部位とのrmsを表示した。第45図において、Aは主鎖原子、Bは側鎖原子、Cは全原子、である。第47図は、1BZF (MODEL 18) における活性部位・リガンド結合解析の結果である。MD計算を0.1nsecまで行ったときの結合解析及び1.0nsecまで行ったとき、また、受容体動的構造クラスタリングを行うときの母集団を100fsecごと及び1000fsecごとで行ったときの結合解析の結果である。評価法は、正解構造との活性部位及びリガンドのrmsで行った。第46図にリガンドドッキングにおける空間点指定のパラメータ値を示した。第46図は、1LUD (MODEL 4) より得られた構造活性相関情報である。

第48図～第50図にタンパク/リガンド複合体と正解構造との比較を示した。第48図は、0～1.0nsecの範囲内で100fsecごとの母集団により作られたクラスターを用いて行った活性部位・リガンド結合である。緑色Aは正解構造の1LUD (MODEL 4)、青色Bは初期構造の1BZF (MODEL 18)、赤色Cはリガンド結合における最適構造、要素色Dはリガンドの正解構

造、水色Eは計算結果によるリガンド、である。リガンドのrmsは0.9614であった。第50図のリガンドを結合させることにより活性部位の主鎖原子のrmsで0.2791の誘導が生じた。第49図において、黒色は正解構造(1ludのmodel 4)を表し、灰色は初期構造(1bzfのmodel 18)を表し、白色は最適構造を表す。第50図は、2,4-diamino-5-(3,4,5-trimethoxy-benzyl)-pyrimidin-1-iumである。1LUDのリガンドである。

Case 2)

heat shock protein 90 (HSP90)である1YERと1YETは、ホモロジーが100.0%でリガンド結合の有無により活性部位の形が異なる。そこで、リガンド結合していない1YERを初期構造に選び、リガンドとしてgeldanamycinを用い、正解構造である1YETと比較することで検証した(第51図)。第51図は、heat shock protein 90の立体構造である。1YETの受容体(緑色A)とリガンド(赤色B)、1YERの受容体(青色C)、である。

第52図に1YERの基準振動解析の結果を示した。二面角 ϕ (橙色A)、 ϕ (緑色B)の揺らぎの大きさを示した。揺らぎが0.0に近いほど分子動力学法(MD)計算において二面角拘束が強くなる。また、STRIDEによる二次構造判定で α -ヘリックス(赤色D)、 β -シート(青色D)を表示した。紫色Cは活性部位である。

第53図、第55図に1YERを用いた拘束二面角分子動力学の結果を示した。正解構造1YETの活性部位とのrmsを表示した。Aは主鎖原子、Bは側鎖原子、Cは全原子、である。第55図は、1YERにおける活性部位・リガンド結合解析の結果である。MD計算を0.1nsecまで行ったときの結合解析及び1.0nsecまで行ったとき、また、受容体動的構造クラスタリングを行うときの母集団を100fsecごと及び1000fsecごとで行ったときの結合解析である。評価法は、正解構造との活性部位及びリガンドのrmsで行った。

第54図にリガンドドッキングにおける空間点指定のパラメータ値を示した。第54図は、1YETより得られた構造活性相関情報である。

第56図および第57図にタンパク／リガンド複合体と正解構造との比較を示した。第56図および第57図は1YER・リガンド結合である。第56図は、
5 0～0.1nsecの範囲内で100fsecごとの母集団により作られたクラスターを用いて行った活性部位・リガンド結合である。緑色Aは正解構造の1YET、青色Bは初期構造の1YER、赤色Cはリガンド結合における最適構造、要素色Dはリガンドの正解構造、水色Eは計算結果によるリガンド、である。リ
10 ガンドのrmsは1.2081である。第57図のリガンドを結合させることにより活性部位の主鎖原子のrmsで0.1619の誘導が生じた。第57図は、geldanamycinである。1YETのリガンドである。

Case 3)

mitogen-activated protein kinase (MAP kinase) である1A9Uと1OUKはホモロジー100.0%でかつ
15 結合しているリガンドが異なることにより活性部位の形が異なる。そこで、1A9Uを初期構造に選び、リガンドとして1OUK中に含まれるリガンドを用い、正解構造である1OUKと比較することで検証した(第58図)。第58図はmitogen-activated protein kinaseの立体構造である。第58図において、1OUKの受容体(緑色A)とリガンド(赤色B)
20)、1A9Uの受容体(青色C)とリガンド(水色D)、である。

第59図に基準振動解析結果を示した。二面角 ϕ (橙色A)、 ϕ (緑色B)の揺らぎの大きさを示した。揺らぎが0.0に近いほど分子動力学法(MD)計算において二面角拘束が強くなる。また、STRIDEによる二次構造判定で α -ヘリックス(赤色D)、 β -シート(青色D)を表示した。紫色Cは活性部位で
25 ある。

第60図、第62図に1YERを用いた拘束二面角分子動力学の結果を示した。第60図は1A9UのMD計算の結果である。第60図に正解構造1OUKの活

性部位との *rms* を表示した。Aは主鎖原子、Bは側鎖原子、Cは全原子、である。第62図は1A9Uにおける活性部位・リガンド結合解析の結果である。MD計算を0.1 *nsec* まで行ったときの結合解析及び1.0 *nsec* まで行ったとき、また、受容体動的構造クラスタリングを行うときの母集団を100 *fs* *sec* ごと及び1000 *fs* *sec* ごとで行ったときの結合解析の結果である。評価法は、正解構造との活性部位及びリガンドの *rms* で行った。

第61図にリガンドドッキングにおける空間点指定のパラメータ値を示した。第61図は1OUKより得られた構造活性相関情報である。

第63図～第65図にタンパク／リガンド複合体と正解構造との比較を示した。第63図～第65図は1A9U・リガンド結合である。第63図は、0～0.1 *nsec* の範囲内で100 *fs* *sec* ごとの母集団により作られたクラスターを用いて行った活性部位・リガンド結合である。緑色Aは正解構造の1OUK、青色Bは初期構造の1A9U、赤色Cはリガンド結合における最適構造、要素色Dはリガンドの正解構造、水色Eは計算結果によるリガンド、である。リガンドの *rms* は1.6112であった。第65図のリガンドを結合させることにより活性部位の主鎖原子の *rms* で0.1871の誘導が生じた。第64図において、黒色は正解構造 (1ouk)、灰色は初期構造 (1a9u)、白色は最適構造、である。第65図は1OUKのリガンドである。4-[5-[2-(1-phenyl-ethylamino)-pyrimidin-4-yl]-1-methyl-4-(3-trifluoromethylphenyl)-1H-imidazol-2-yl]-piperidineである。

Case 1) ～Case 3) に示す通り、リガンド探索装置100により作成されるタンパク質／リガンド複合体モデルは、誘導結合型のタンパク質／リガンド複合体の立体構造を精度よく予測可能であることが分った。

25 [実施例4]

(Fxaを用いた *in silico* Screeningへの応用例)

上述した実施形態のリガンド探索装置100により、セリンプロテアーゼの1

種である F x a の立体構造 (第 6 6 図) を用い、化合物データベースから F x a に結合する可能性のあるリガンドを探索した。立体構造には 1 A I X を用い、リガンドデータベースとして PDB データベースより収集した 3 6 3 3 種類のリガンドを用いた。上述した実施形態に従い、in silico screening を行った。その結果を第 6 7 図に示した。

第 6 7 図は、化合物データベース中のリガンドのうち、1 A I X との相互作用エネルギーの上位 1 0 0 個を示している。第 6 7 図において、太字は 1 A I X 中に含まれているリガンドであり、斜線はセリンプロテアーゼである。PDB code とは、リガンドが含まれているものの PDB code を示す。第 6 7 図には、1 A I X にもともと含まれているリガンドがランキング 1 9 位に入っている。

ランキング 1 9 位におけるタンパク質／リガンド複合体構造を第 6 8 図および第 6 9 図に示した。第 6 8 図において、白色は受容体、黒色は 1 A I X のリガンドである。第 6 9 図は 1 A I X 中のリガンドである。

第 6 7 図中のランキング 3 5 位、3 8 位、8 0 位はすべてセリンプロテアーゼに結合するリガンドである。

これらの構造とタンパク質／リガンド複合体構造を、第 7 0 図および第 7 1 図、第 7 2 図および第 7 3 図、第 7 4 図および第 7 5 図に示した。第 7 0 図および第 7 1 図はランキング 3 5 位におけるタンパク質／リガンド複合体構造である。第 7 0 図において、白色は受容体、黒色は 1 A U J のリガンドである。第 7 1 図は 1 A U J 中のリガンドである。第 7 2 図および第 7 3 図はランキング 3 8 位におけるタンパク質／リガンド複合体構造である。第 7 2 図において、白色は受容体、黒色は正解 (1 F O R) のリガンドであり、RMS は 1. 5 0 0 であった。第 7 3 図は 1 F O R 中のリガンドである。第 7 4 図および第 7 5 図はランキング 8 0 位におけるタンパク質／リガンド複合体構造である。第 7 4 図において、白色は受容体、黒色は 1 K 1 M のリガンドである。第 7 5 図は 1 K 1 M 中のリガンドである。

これらの結果から、本発明により、化合物データベースからもっともらしい化

合物を選択することが可能であることが分かった。

[実施例 5]

(異なる条件での *in silico* スクリーニング)

5 上述した実施形態のリガンド探索装置 100 により、構造活性相関 (SAR) の情報により順位が変動することを検証した。また、受容体を固定した場合の順位の変動も検証した。

ここでは、*severe acute respiratory syndrome* (SARS) のプロテアーゼを用いた *in silico screening* を行った。初期構造にはリガンドを含まない 1UK3 (B鎖) を、またリ
10 ガンドを含む 1UK4 (B鎖) のリガンド結合様式を構造活性相関情報として利用した。活性部位は 1UK4 (B鎖) のリガンドの各原子から半径 10 Å 以内に含まれる受容体残基部位である。リガンドデータベースとして、PDB より収集した 3633 種類のリガンドを用いた。ただし、結合解析で利用する受容体動的構造クラスターには 0 ~ 0.1 nsec の範囲内で 100 fsec ごとの母集団
15 で作られたものを使用した。また、水素原子を除いて計算した。

第 76 図は SARS プロテアーゼの立体構造である。1UK4 (B鎖) の受容体 (緑色 A) とリガンド (赤色 B)、1UK3 (B鎖) の受容体 (青色 C) が示されている。

第 77 図に 1UK3 (B鎖) の基準振動解析の結果を示した。二面角 ϕ (橙色
20 A)、 ϕ (緑色 B) の揺らぎの大きさを示した。揺らぎが 0.0 に近いほど分子動力学法 (MD) 計算において二面角拘束が強くなる。また、STRIDE による二次構造判定で α -ヘリックス (赤色 D)、 β -シート (青色 D) を表示した。紫色 C は活性部位である。

第 78 図に 1UK3 の分子動力学計算の結果を示した。第 78 図は 1UK3 (B鎖) の MD。1UK4 (B鎖) の活性部位との rms を表示した。A は主鎖原子、B は側鎖原子、C は全原子、である。

Case 1) SAR 4 カ所指定

第79図に1UK4の活性部以内での空間指定を示した。第79図は1UK4 (B鎖) より得られた構造活性相関情報である。第80図に、1UK3 (B鎖) における *in silico* スクリーニングの結果を示した。

第81図には、1UK3と1UK4とに関し、正解構造との比較を示した。順位は25位である。緑色Aは1UK4 (B鎖)、青色Bは初期構造の1UK3 (B鎖)、赤色Cはリガンド結合における最適構造、要素色Dは1UK4のペプチド性リガンド (ASN-SER-THR-LEU-GLN) の正解構造、水色Eは計算結果によるリガンド、である。リガンドの *rms* は2.5721であった。正解構造との活性部位の主鎖原子の *rms* は、初期構造では1.0248、最適構造では1.0792、であった。

第82図および第83図に、*in silico* スクリーニングの順位1を示した。第83図には1QF4のリガンド (C8-R) -hydantocidin 5'-phosphateを示した。

Case 2) SAR3カ所指定

第84図に1UK3の活性部位内での空間指定を示した。第84図は1UK3 (B鎖) より得られた構造活性相関情報である。

第85図には、1UK3 (B鎖) と1UK4 (B鎖) との比較結果であり、1UK3での最適構造と正解構造との比較を示した。順位は49である。緑色Aは1UK4 (B鎖)、青色Bは初期構造の1UK3 (B鎖)、赤色Cはリガンド結合における最適構造、要素色Dは1UK4のペプチド性リガンド (ASN-SER-THR-LEU-GLN) の正解構造、水色Eは計算結果によるリガンドである。リガンドの *rms* は2.0057であった。正解構造との活性部位の主鎖原子の *rms* は、初期構造では1.0248、最適構造では1.0469であった。

第86図には、SAR3カ所指定で実行した *in silico* スクリーニングの結果を示した。

Case 3) SAR5カ所指定

第87図には1UK3の活性部位内での空間指定を示した。第87図は1UK3 (B鎖) より得られた構造活性相関情報に関する図である。第88図には、SAR5カ所指定で実行した *in silico* スクリーニング (ハイスループットスクリーニング) 結果を示した。

- 5 第89図には、1UK3での最適構造と正解構造との比較を示した。第89図は1UK3 (B鎖) と1UK4 (B鎖) との比較結果である。順位は2位である。緑色Aは1UK4 (B鎖)、青色Bは初期構造の1UK3 (B鎖)、赤色Cはリガンド結合における最適構造、要素色Dは1UK4のペプチド性リガンド (ASN-SER-THR-LEU-GLN) の正解構造、水色Eは計算結果によるリ
- 10 ガンド、である。リガンドのrmsは1.2578であった。正解構造との活性部位の主鎖原子のrmsは、初期構造では1.0248、最適構造では1.1620であった。

Case 4) リガンド原子タイプ指定の変更

- 第90図に1UK3の活性部位内での空間指定を示した。第90図は1UK3 (B鎖) より得られた構造活性相関情報に関する図である。第91図には、リガンド原子タイプ指定変更で実行した *in silico* スクリーニング (ハイスループットスクリーニング) の結果を示した。
- 15

- 第92図には、1UK3での最適構造と正解構造との比較を示した。第92図は1UK3 (B鎖) と1UK4 (B鎖) との比較結果である。順位は774位である。緑色Aは1UK4 (B鎖)、青色Bは初期構造の1UK3 (B鎖)、赤色Cはリガンド結合における最適構造、要素色Dは1UK4のペプチド性リガンド (ASN-SER-THR-LEU-GLN) の正解構造、水色Eは計算結果によるリガンド、である。リガンドのrmsは2.5216であった。正解構造との活性部位の主鎖原子のrmsは、初期構造では1.0248、最適構造では1.0792であった。
- 20
- 25

Case 5) 受容体固定

第93図に活性部位内での空間指定を示した。第93図は1UK4 (B鎖) よ

り得られた構造活性相関情報に関する図である。第94図に、受容体を固定した状態で実行した *in silico* スクリーニング (ハイスループットスクリーニング) の結果を示した。

第95図には、1UK3と1UK4を重ね合わせたリガンドと計算結果のリガンドとの比較を示した。順位は39位である。灰色は1UK3の活性部位構造、黒色は1UK3と1UK4を重ね合わせたリガンド、白色は計算結果のリガンド、である。

Case 1) ~ Case 4) を見ると、SARの指定が多いほど、参考にしたリガンドの順位が良くなった。つまり、参考にできるリガンドの結合情報が信頼できる場合にはSARの情報を多くした *in silico* スクリーニングを行い、信頼性に欠ける場合にはSARの情報数を減らし、さらにリガンド原子タイプ指定の幅を広げることで、様々なリガンドがランキング上位に分布した。そして、その分布情報をもとにSAR情報を作り変えて *in silico* スクリーニングを実行するとより信頼性の持てる結果が出力された。

Case 1) と Case 5) を見ると、受容体の動的構造の有無による順位変動を示している。これは、リガンドの動きのみの最適化に比べ、リガンド及び受容体それぞれが動く最適化の方が原子のぶつかりをさけることに優れている。従って、同じ位置に配置するための最適化エネルギーに差が生じた。

[実施例6]

(二面角拘束分子動力学計算のパラメータに関するMDパラメータの分布)

FMN-binding proteinにおける二面角拘束MDパラメータの分布に関し説明する。ここでは、上述した実施形態のリガンド探索装置100により、1LUD以外のNMR構造でも二面角拘束分子動力学計算及びクラスタリングのパラメータが同様の結果を生じるのかを検証した。そこで、FMN-binding proteinのNMR構造(PDB code: 1AXJ)のMODEL 1を初期構造に選んだ。評価法は、受容体動的構造クラスタリングのパラメーター ($\alpha = 80.0\%$ 、 $\beta = 0.4 \text{ \AA}$) を固定したこと以外は実施例1

に従った。

第96図に1AXJにおける二面角拘束分子動力学計算のパラメータ決定のスコアの分布状況を示した。第96図は1AXJにおける二面角拘束MDパラメータの分布である。Aに近い部分ほどスコアで小さい。1LUDの時と同様に二面角拘束の最大値800、最小値0では良い結果を示した。

[実施例7]

(二面角拘束MD)

ここでは、上述した実施形態のリガンド探索装置100により、主鎖二面角拘束MDで各原子の動的構造を検証した。なお、時として基準振動解析が収束せず二面角揺らぎ情報が得られないことがある。そこで、第13図により主鎖二面角に対して均一な拘束(500)でMDを行ったときでも良い結果になっていることから、この場合における動的構造も検証した。実施例1に従い、拘束なし、二面角揺らぎを用いた拘束及び均一な拘束(500)の条件のもとMDを行った。

第97図～第108図に1LUDに対して行った分子動力学計算の各原子における動的挙動の結果を示した。第97図～第108図は1LUD(MODEL1)のMD計算の結果である。第97図および第98図は活性部位の主鎖原子、第99図および第100図は受容体の主鎖原子、第101図および第102図は活性部位の側鎖原子、第103図および第104図は受容体の側鎖原子、第105図および第106図は活性部位の全原子、第107図および第108図は受容体の全原子、における動的挙動の結果を示す図である。1LUDのPDBファイル内に記載されている24種類の各モデル構造をMODEL1と活性部位の主鎖原子においてrmsを計算し、その平均rmsを点線で表示した。二面角拘束があるとき(A)とないとき(B)及び二面角拘束が500で一定(C)において、活性部位の主鎖原子の初期構造からのずれをrmsで表示した。

ここには記載しないが、1CBQ、1J9G、1MMB、1BZF(MODEL18)、1YER、1A9U及び1UK3(B鎖)に関しても主鎖二面角揺らぎに基づく拘束MDの結果を見ると、第109図～第111図と同様に主鎖原子

の抑制があると、拘束のない側鎖原子にも一定動きを示した。よって、受容体の動きにおいて主鎖原子の動きの比重が大きいことが理解できた。

[実施例 8]

(異なる条件での結合解析)

- 5 ここでは、上述した実施形態のリガンド探索装置 100 により、二面角拘束 MD 及びクラスタリングのパラメータが異なっても誘導が生じることを検証した。拘束の最大値を 100 に、最小値を 0 に設定し、受容体動的構造クラスタリング定数を $\alpha = 80.0\%$ および $\beta = 1.0 \text{ \AA}$ に設定して、その他は実施例 2 に従った。ただし、受容体動的構造クラスターには $0 \sim 0.1 \text{ nsec}$ の範囲内で 10
- 10 0 fsec ごとの母集団で作られたものを使用した。また、活性部位の定義は、リガンドの各原子から半径 6 Å 以内にある受容体残基とした。

第 109 図～第 111 図には、異なる条件で受容体／リガンド結合の結果を示した。

- (i) 1BZF (MODEL 18) で、リガンド結合により活性部位の主鎖原子
- 15 rms で 0.2686 の誘導が生じた (第 109 図)。活性部位全体の rms では 0.1224 の誘導が生じた。リガンドの rms は 0.8526 であった。

(ii) 1YER で、リガンド結合により活性部位の主鎖原子 rms で 0.2376 の誘導が生じた (第 110 図)。活性部位全体の rms では 0.0816 の誘導が生じた。リガンドの rms は 0.7246 であった。

- 20 (iii) 1A9U で、リガンド結合により活性部位の主鎖原子 rms で 0.2150 の誘導が生じた (第 111 図)。活性部位全体の rms では 0.0464 の誘導が生じた。リガンドの rms は 0.9464 であった。

ただし、緑色は正解構造、青色は初期構造、赤色は最適構造。要素色は正解リガンド、水色は最適リガンド、である。

- 25 第 109 図～第 111 図で示すように、各条件が異なっているにもかかわらず、与えられた条件の中で最適な結果を生じることができた。

[実施例 9]

(正解構造を初期構造に選んだときの結合解析)

ここでは、DHFRの1BZF及び1LUDはリガンドの結合様式が似ているので、上述した実施形態のリガンド探索装置100により、構造活性相関情報を一部変更し1BZFのリガンドの結合解析を行った。条件としては、初期構造に
5 1BZF (MODEL18) を使用し、0～0.1nsecまでの母集団より作成したクラスターを使用した。

第112図には、1BZFの活性部位内における空間指定を示した。第112図は1BZF用に変更した構造活性相関情報に関する図である。第113図および第114図には、1BZFにおける受容体／リガンド結合の結果を示した。第
10 113図および第114図は1BZF (MODEL18) のリガンド結合解析の結果である。

(i)最適化したときの受容体には初期構造を選択した。リガンドのrmsは0.8884であった(第113図)。

(ii) trimetrexate、1BZF (MODEL18) のリガンドで
15 ある(第114図)。

初期構造が元々PDBに登録されていた構造、つまり最適構造であったため、計算結果でもそれが第113図および第114図のように再現できた。

産業上の利用可能性

20 以上のように、本発明にかかるリガンド探索装置、リガンド探索方法、プログラム、および記録媒体は、医農薬の分子設計等を中心に、受容体／リガンド結合の解析を行う分野(医薬品設計)において、極めて有用であると考えられる。本発明は、産業上多くの分野、特に医薬品、食品、化粧品、医療、構造解析、機能解析等の分野で広く実施することができ、故に極めて有用である。

請求の範囲

1. 単数または複数鎖のタンパク質の座標データが与えられた場合に、当該タンパク質と結合するリガンドを探索するリガンド探索装置において、

- 5 上記タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択する構造変化後タンパク質座標データ選択手段と、

上記構造変化後タンパク質座標データ選択手段にて選択された上記構造変化後タンパク質座標データから、上記リガンドと重ね合わせを行う空間点を指定する

- 10 空間点指定手段と、

上記空間点指定手段にて指定された上記空間点と、上記リガンドのリガンド座標データとを用いて、上記タンパク質と上記リガンドとが結合した場合の相互作用関数を計算する相互作用関数計算手段と、

- 15 上記相互作用関数計算手段により計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドを評価するリガンド評価手段と、

を備えたことを特徴とするリガンド探索装置。

2. 請求の範囲第1項に記載のリガンド探索装置において、上記相互作用関数計算手段は、以下の数式1に示す $Sscore(i, j)$ を用いて上記相互作用

- 20 関数を計算することを特徴とするリガンド探索装置。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp\{-(d_y^s - d_y^c)^2\} - \beta \right] / \frac{(d_y^s + d_y^c)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases}$$

... (数式1)

(ここで、 d_{ij}^a は該当タンパク質中の i 番目と j 番目の空間点距離である。
また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。また、 α は、該
当タンパク質中の空間点群と化合物が完全に重なりあった場合に $S.s.c.o.r.e$ (
5 i, j) を最大値とするための定数である。また、 β は重なりと定義できる限界
値を与えるための定数である。)

3. 請求の範囲第1項または第2項に記載のリガンド探索装置において、上記
相互作用関数計算手段は、
10 上記相互作用関数のスコアが最大になるように最適化する相互作用関数最適化
手段、
をさらに備えたことを特徴とするリガンド探索装置。

4. 請求の範囲第3項に記載のリガンド探索装置において、上記相互作用関数
15 計算手段は、
上記相互作用関数最適化手段により上記相互作用関数を最適化した後に、重ね
あわせた上記リガンドに対して、上記タンパク質との相互作用エネルギーを計算
し、当該相互作用エネルギーについてリガンド立体構造データのコンフォメーシ
ョンを微調整しながら最適化する相互作用エネルギー最適化手段、
20 をさらに備えたことを特徴とするリガンド探索装置。

5. 請求の範囲第4項に記載のリガンド探索装置において、上記リガンド評価
手段は、
上記相互作用エネルギー最適化手段により最適化した後に、上記リガンド立体
25 構造データのコンフォメーションを大きく変動させた後、再度、上記相互作用関
数計算手段を実行し、上記相互作用関数計算手段により計算された上記相互作用
関数に基づいて当該タンパク質と結合する上記リガンドの再評価を行う再評価手

段、

をさらに備えたことを特徴とするリガンド探索装置。

6. 請求の範囲第1項から第5項のいずれか1つに記載のリガンド探索装置において、上記構造変化後タンパク質座標データ選択手段は、

上記誘導適合パラメータおよび／または上記構造変化後タンパク質座標データを算出するときに、上記タンパク質座標データに対し基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、当該ゆらぎの大きさを拘束条件として、分子動力学計算を行うこと、

- 10 を特徴とするリガンド探索装置。

7. 請求の範囲第6項に記載のリガンド探索装置において、上記構造変化後タンパク質座標データ選択手段は、

- 15 基準振動計算より主鎖原子の2面角のゆらぎの値を算出し、当該ゆらぎ値を以下の数式2または数式3に示す分子動力学計算における力の定数Kとすることにより、分子動力学計算を行うこと、

を特徴とするリガンド探索装置。

$$E_{rot} = K_{rot} (\phi - \phi_0)^2 \quad \dots \text{(数式2)}$$

20

(E_{rot} はタンパク質の立体構造中において主鎖原子の2面角のエネルギーを示す。 ϕ は主鎖原子の2面角である。 ϕ_0 は主鎖原子の2面角の標準値である。ここで、 K_{rot} の値が大きい場合は、 ϕ は ϕ_0 に拘束される。)

$$25 \quad E_{pos} = K_{pos} (r - r_0)^2 \quad \dots \text{(数式3)}$$

(E_{pos} はタンパク質の立体構造中において主鎖原子の位置のエネルギーを

示す。r は主鎖原子の座標である。r₀ は主鎖原子の座標の標準値である。ここで、K_{p o s} の値が大きい場合は、r は r₀ に拘束される。)

8. 請求の範囲第1項から第7項のいずれか1つに記載のリガンド探索装置において、上記相互作用関数計算手段は、

上記相互作用関数として、相互作用エネルギー関数にタンパク質の動的性質を表現する動的性質関数を「弾性エネルギー」として加えて用いること、
を特徴とするリガンド探索装置。

9. 請求の範囲第8項に記載のリガンド探索装置において、上記相互作用関数計算手段は、

「弾性エネルギー」として、タンパク質の局所的な柔らかさを考慮し、以下の数式4に示す関数「U_{衝突}」を適応すること、
を特徴とするリガンド探索装置。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2 \quad \dots \text{(数式4)}$$

(Mは衝突を不許可とする活性部位の原子の数であり、Nはリガンドの原子の数である。活性部位における動的挙動の少ない側鎖原子及び主鎖原子のみのi番目の原子とリガンドのj番目の原子との原子間距離Rが衝突距離「R_{衝突}(i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。)

10. 請求の範囲第1項から第7項のいずれか1つに記載のリガンド探索装置において、上記相互作用関数計算手段は、

上記相互作用関数として、相互作用エネルギー関数に、当該タンパク質の基準振動解析結果または二次構造判定結果をタンパク質の動的性質を表現する動的性

質関数として加えて用いること、
を特徴とするリガンド探索装置。

1 1. 単数または複数鎖のタンパク質の座標データが与えられた場合に、当該
5 タンパク質と結合するリガンドを探索するリガンド探索方法において、

上記タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択する構造変化後タンパク質座標データ選択ステップと、

上記構造変化後タンパク質座標データ選択ステップにて選択された上記構造変化後タンパク質座標データから、上記リガンドと重ね合わせを行う空間点を指定する空間点指定ステップと、

上記空間点指定ステップにて指定された上記空間点と、上記リガンドのリガンド座標データとを用いて、上記タンパク質と上記リガンドとが結合した場合の相互作用関数を計算する相互作用関数計算ステップと、

15 上記相互作用関数計算ステップにより計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドを評価するリガンド評価ステップと、
を含むことを特徴とするリガンド探索方法。

1 2. 請求の範囲第 1 1 項に記載のリガンド探索方法において、上記相互作用関数計算ステップは、以下の数式 1 に示す $S\ score(i, j)$ を用いて上記相互作用関数を計算することを特徴とするリガンド探索方法。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp\{-(d_{ij}^s - d_{ij}^c)^2\} - \beta \right] / \frac{(d_{ij}^s + d_{ij}^c)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases} \quad \dots \text{ (数式 1)}$$

(ここで、 d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。
 10 また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。また、 α は、該
 該当タンパク質中の空間点群と化合物が完全に重なりあった場合に $Sscore(i, j)$ を
 最大値とするための定数である。また、 β は重なりと定義できる限界
 値を与えるための定数である。)

15 13. 請求の範囲第11項または第12項に記載のリガンド探索方法において、
 上記相互作用関数計算ステップは、

上記相互作用関数のスコアが最大になるように最適化する相互作用関数最適化
 ステップ、

をさらに含むことを特徴とするリガンド探索方法。

20

14. 請求の範囲第13項に記載のリガンド探索方法において、上記相互作用
 関数計算ステップは、

上記相互作用関数最適化ステップにより上記相互作用関数を最適化した後に、
 重ねあわせた上記リガンドに対して、上記タンパク質との相互作用エネルギーを
 25 計算し、当該相互作用エネルギーについてリガンド立体構造データのコンフォメ
 ーションを微調整しながら最適化する相互作用エネルギー最適化ステップ、
 をさらに含むことを特徴とするリガンド探索方法。

15. 請求の範囲第14項に記載のリガンド探索方法において、上記リガンド評価ステップは、

5 上記相互作用エネルギー最適化ステップにより最適化した後に、上記リガンド立体構造データのコンフォメーションを大きく変動させた後、再度、上記相互作用関数計算ステップを実行し、上記相互作用関数計算ステップにより計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドの再評価を行う再評価ステップ、

をさらに含むことを特徴とするリガンド探索方法。

10

16. 請求の範囲第11項から第15項のいずれか1つに記載のリガンド探索方法において、上記構造変化後タンパク質座標データ選択ステップは、

15 上記誘導適合パラメータおよび/または上記構造変化後タンパク質座標データを算出するときに、上記タンパク質座標データに対し基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、当該ゆらぎの大きさを拘束条件として、分子動力学計算を行うこと、

を特徴とするリガンド探索方法。

20 17. 請求の範囲第16項に記載のリガンド探索方法において、上記構造変化後タンパク質座標データ選択ステップは、

基準振動計算より主鎖原子の2面角のゆらぎの値を算出し、当該ゆらぎ値を以下の数式2または数式3に示す分子動力学計算における力の定数Kとすることにより、分子動力学計算を行うこと、

を特徴とするリガンド探索方法。

25

$$E_{rot} = K_{rot} (\phi - \phi_0)^2 \quad \dots \quad \text{(数式2)}$$

(E_{rot} はタンパク質の立体構造中において主鎖原子の2面角のエネルギーを示す。 ϕ は主鎖原子の2面角である。 ϕ_0 は主鎖原子の2面角の標準値である。ここで、 K_{rot} の値が大きい場合は、 ϕ は ϕ_0 に拘束される。)

$$5 \quad E_{pos} = K_{pos} (r - r_0)^2 \quad \dots \text{(数式3)}$$

(E_{pos} はタンパク質の立体構造中において主鎖原子の位置のエネルギーを示す。 r は主鎖原子の座標である。 r_0 は主鎖原子の座標の標準値である。ここで、 K_{pos} の値が大きい場合は、 r は r_0 に拘束される。)

10

18. 請求の範囲第11項から第17項のいずれか1つに記載のリガンド探索方法において、上記相互作用関数計算ステップは、

上記相互作用関数として、相互作用エネルギー関数にタンパク質の動的性質を表現する動的性質関数を「弾性エネルギー」として加えて用いること、

15

を特徴とするリガンド探索方法。

19. 請求の範囲第18項に記載のリガンド探索方法において、上記相互作用関数計算ステップは、

「弾性エネルギー」として、タンパク質の局所的な柔らかさを考慮し、以下の

20

数式4に示す関数「U衝突」を適応すること、

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

25

... (数式4)

(M は衝突を不許可とする活性部位の原子の数であり、 N はリガンドの原子の数である。活性部位における動的挙動の少ない側鎖原子及び主鎖原子のみの i 番

目の原子とリガンドの j 番目の原子との原子間距離 R が衝突距離「 R 衝突 (i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。)

20. 請求の範囲第11から第17項のいずれか1つに記載のリガンド探索方法において、上記相互作用関数計算ステップは、

上記相互作用関数として、相互作用エネルギー関数に、当該タンパク質の基準振動解析結果または二次構造判定結果をタンパク質の動的性質を表現する動的性質関数として加えて用いること、

を特徴とするリガンド探索方法。

10

21. 単数または複数鎖のタンパク質の座標データが与えられた場合に、当該タンパク質と結合するリガンドを探索するリガンド探索方法をコンピュータに実行させるプログラムにおいて、

上記タンパク質の座標データに対して、誘導適合を反映した誘導適合パラメータを用いて動的挙動を考慮した構造変化を行い、構造変化後タンパク質座標データを選択する構造変化後タンパク質座標データ選択ステップと、

上記構造変化後タンパク質座標データ選択ステップにて選択された上記構造変化後タンパク質座標データから、上記リガンドと重ね合わせを行う空間点を指定する空間点指定ステップと、

20 上記空間点指定ステップにて指定された上記空間点と、上記リガンドのリガンド座標データとを用いて、上記タンパク質と上記リガンドとが結合した場合の相互作用関数を計算する相互作用関数計算ステップと、

上記相互作用関数計算ステップにより計算された上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドを評価するリガンド評価ステップと、

25 を含むリガンド探索方法をコンピュータに実行させることを特徴とするプログラム。

22. 請求の範囲第21項に記載のプログラムにおいて、上記相互作用関数計算ステップは、以下の数式1に示す $Sscore(i, j)$ を用いて上記相互作用関数を計算することを特徴とするプログラム。

$$Sscore(i, j) = \sum_j \begin{cases} i \neq j \text{ のとき} \\ \alpha \times \left[\exp\{-(d_{ij}^s - d_{ij}^c)^2\} - \beta \right] / \frac{(d_{ij}^s + d_{ij}^c)^2}{2} \\ i = j \text{ のとき} \\ \alpha \times (1 - \beta) \end{cases}$$

... (数式1)

(ここで、 d_{ij}^s は該当タンパク質中の i 番目と j 番目の空間点距離である。また、 d_{ij}^c は化合物中の i 番目と j 番目の原子間距離である。また、 α は、該当タンパク質中の空間点群と化合物が完全に重なりあった場合に $Sscore(i, j)$ を最大値とするための定数である。また、 β は重なりと定義できる限界値を与えるための定数である。)

23. 請求の範囲第21項または第22項に記載のプログラムにおいて、上記相互作用関数計算ステップは、
 上記相互作用関数のスコアが最大になるように最適化する相互作用関数最適化ステップ、
 をさらに含むことを特徴とするプログラム。

24. 請求の範囲第23項に記載のプログラムにおいて、上記相互作用関数計算ステップは、

上記相互作用関数最適化ステップにより上記相互作用関数を最適化した後に、
 重ねあわせた上記リガンドに対して、上記タンパク質との相互作用エネルギーを

計算し、当該相互作用エネルギーについてリガンド立体構造データのコンフォメーションを微調整しながら最適化する相互作用エネルギー最適化ステップ、
をさらに含むことを特徴とするプログラム。

- 5 25. 請求の範囲第24項に記載のプログラムにおいて、上記リガンド評価ステップは、

上記相互作用エネルギー最適化ステップにより最適化した後に、上記リガンド立体構造データのコンフォメーションを大きく変動させた後、再度、上記相互作用関数計算ステップを実行し、上記相互作用関数計算ステップにより計算された
10 上記相互作用関数に基づいて当該タンパク質と結合する上記リガンドの再評価を行う再評価ステップ、
をさらに含むことを特徴とするプログラム。

26. 請求の範囲第21項から第25項のいずれか1つに記載のプログラムに
15 において、上記構造変化後タンパク質座標データ選択ステップは、

上記誘導適合パラメータおよび／または上記構造変化後タンパク質座標データを算出するときに、上記タンパク質座標データに対し基準振動計算を行い、各アミノ酸のゆらぎの大きさを求め、当該ゆらぎの大きさを拘束条件として、分子動力学計算を行うこと、
20 を特徴とするプログラム。

27. 請求の範囲第26項に記載のプログラムにおいて、上記構造変化後タンパク質座標データ選択ステップは、

基準振動計算より主鎖原子の2面角のゆらぎの値を算出し、当該ゆらぎ値を以下の数式2または数式3に示す分子動力学計算における力の定数Kとすることにより、分子動力学計算を行うこと、
25 を特徴とするプログラム。

$$E_{rot} = K_{rot} (\phi - \phi_0)^2 \quad \dots \text{(数式2)}$$

(E_{rot} はタンパク質の立体構造中において主鎖原子の2面角のエネルギーを示す。 ϕ は主鎖原子の2面角である。 ϕ_0 は主鎖原子の2面角の標準値である。ここで、 K_{rot} の値が大きい場合は、 ϕ は ϕ_0 に拘束される。)

$$E_{pos} = K_{pos} (r - r_0)^2 \quad \dots \text{(数式3)}$$

(E_{pos} はタンパク質の立体構造中において主鎖原子の位置のエネルギーを示す。 r は主鎖原子の座標である。 r_0 は主鎖原子の座標の標準値である。ここで、 K_{pos} の値が大きい場合は、 r は r_0 に拘束される。)

28. 請求の範囲第21項から第27項のいずれか1つに記載のプログラムにおいて、上記相互作用関数計算ステップは、

上記相互作用関数として、相互作用エネルギー関数にタンパク質の動的性質を表現する動的性質関数を「弾性エネルギー」として加えて用いること、
を特徴とするプログラム。

29. 請求の範囲第28項に記載のプログラムにおいて、上記相互作用関数計算ステップは、

「弾性エネルギー」として、タンパク質の局所的な柔らかさを考慮し、以下の数式4に示す関数「U衝突」を適応すること、
を特徴とするプログラム。

$$U_{\text{衝突}} = \sum_{i=1}^M \sum_{j=1}^N \phi(i, j)$$

$$\phi(i, j) = K_{\text{衝突}} * (R_{\text{衝突}}(i, j) - R)^2$$

... (数式4)

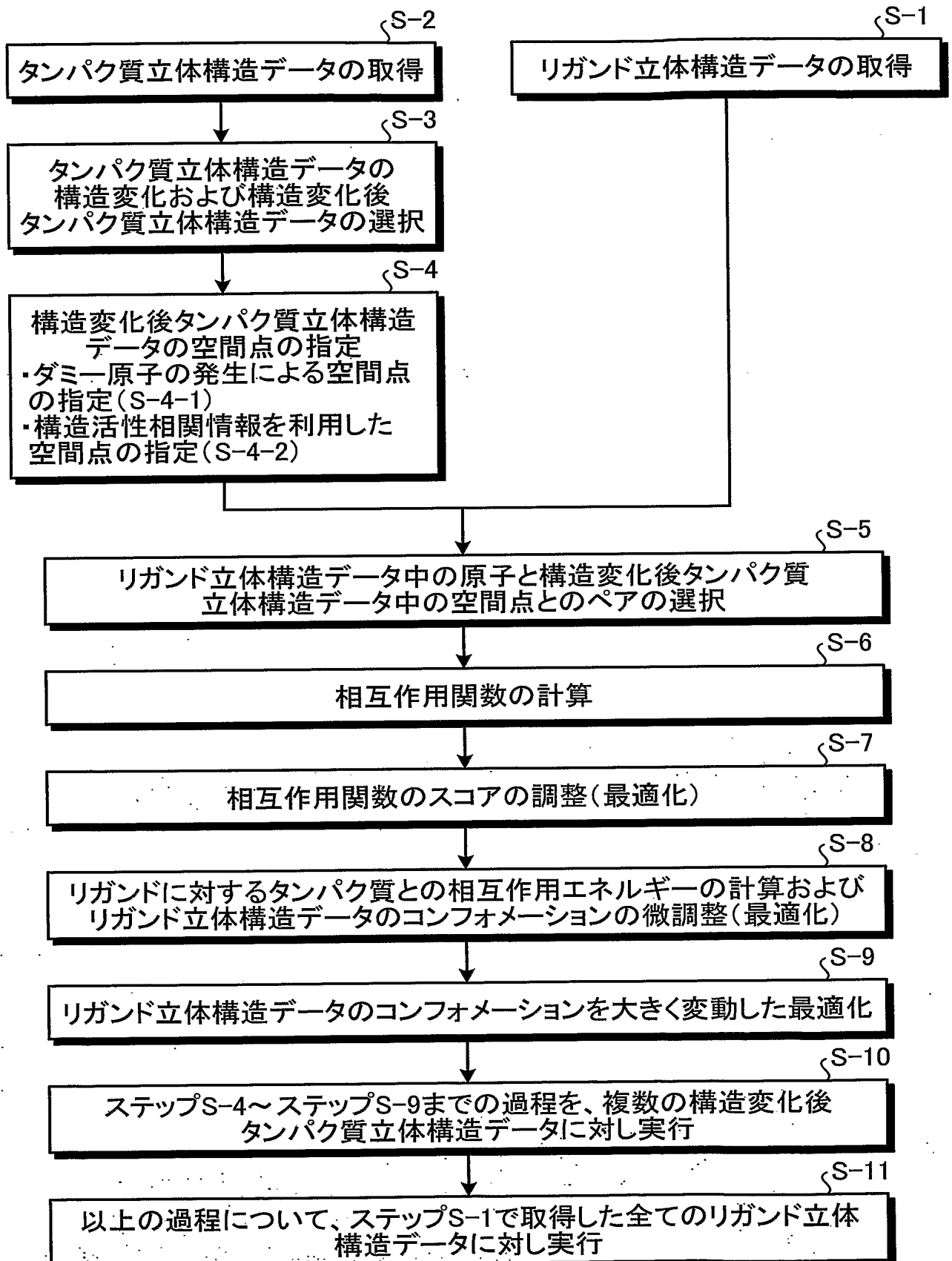
(Mは衝突を不許可とする活性部位の原子の数であり、Nはリガンドの原子の数である。活性部位における動的挙動の少ない側鎖原子及び主鎖原子のみの i 番目の原子とリガンドの j 番目の原子との原子間距離 R が衝突距離「R 衝突 (i, j)」以内のとき、 $\phi(i, j)$ を計算するように定義する。)

30. 請求の範囲第21項から第27項のいずれか1つに記載のプログラムにおいて、上記相互作用関数計算ステップは、

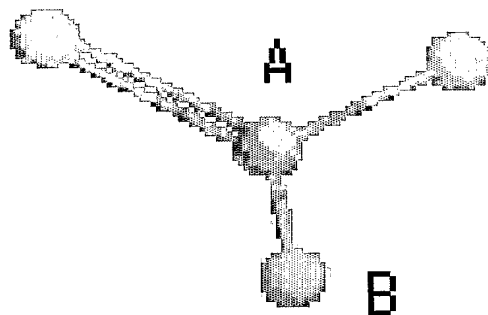
上記相互作用関数として、相互作用エネルギー関数に、当該タンパク質の基準振動解析結果または二次構造判定結果をタンパク質の動的性質を表現する動的性質関数として加えて用いること、
を特徴とするプログラム。

31. 請求の範囲第21項から第30項のいずれか1つに記載のプログラムを記録したことを特徴とするコンピュータ読み取り可能な記録媒体。

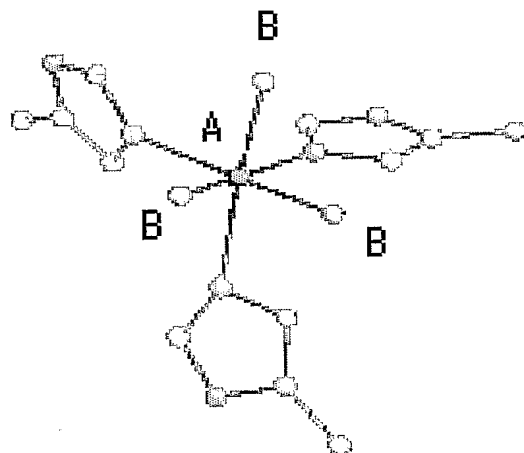
第1図



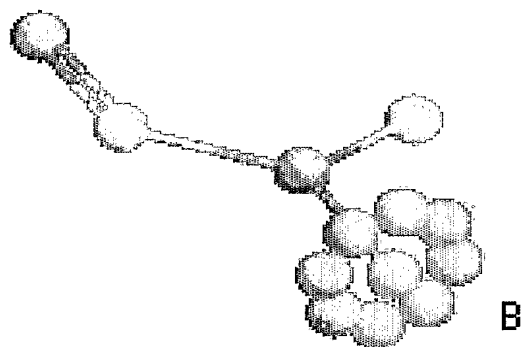
第2図



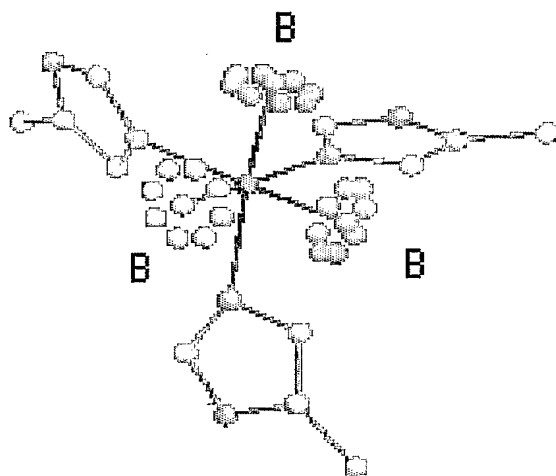
第3図



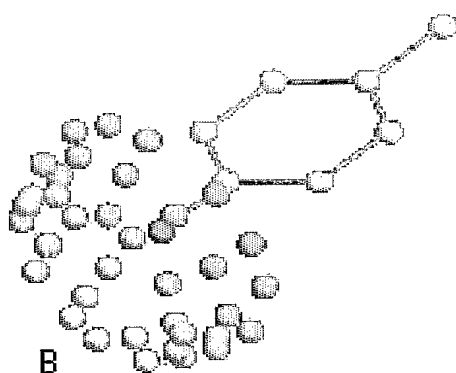
第4図



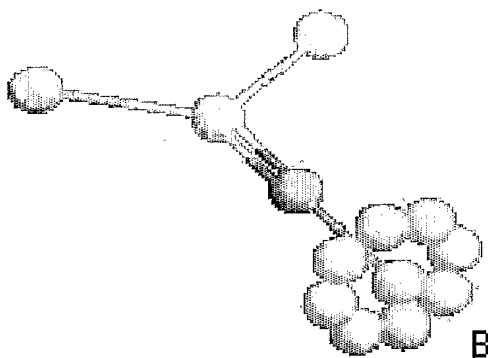
第5図



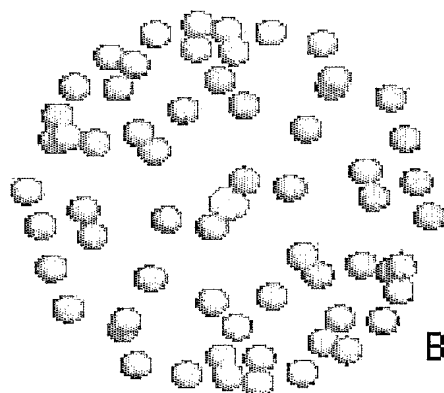
第6図



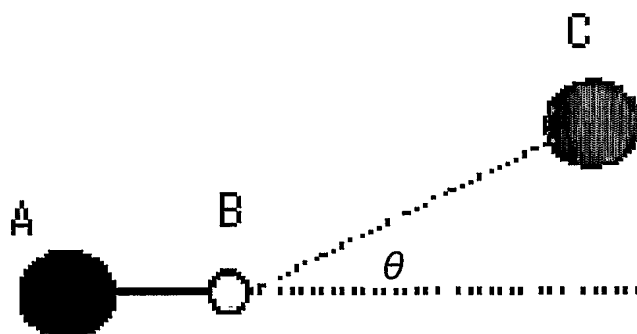
第7図



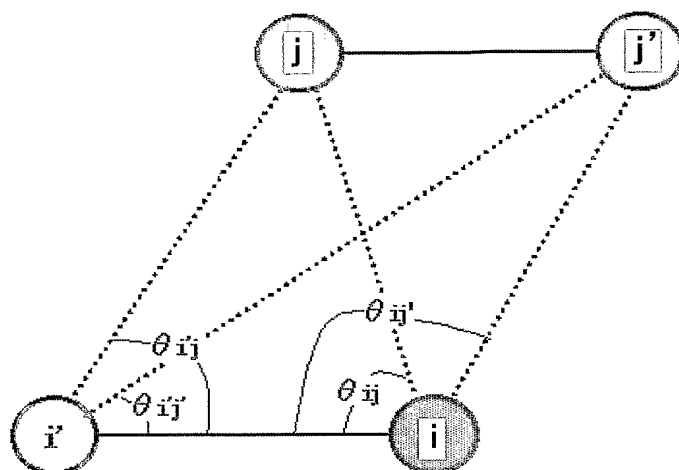
第8図



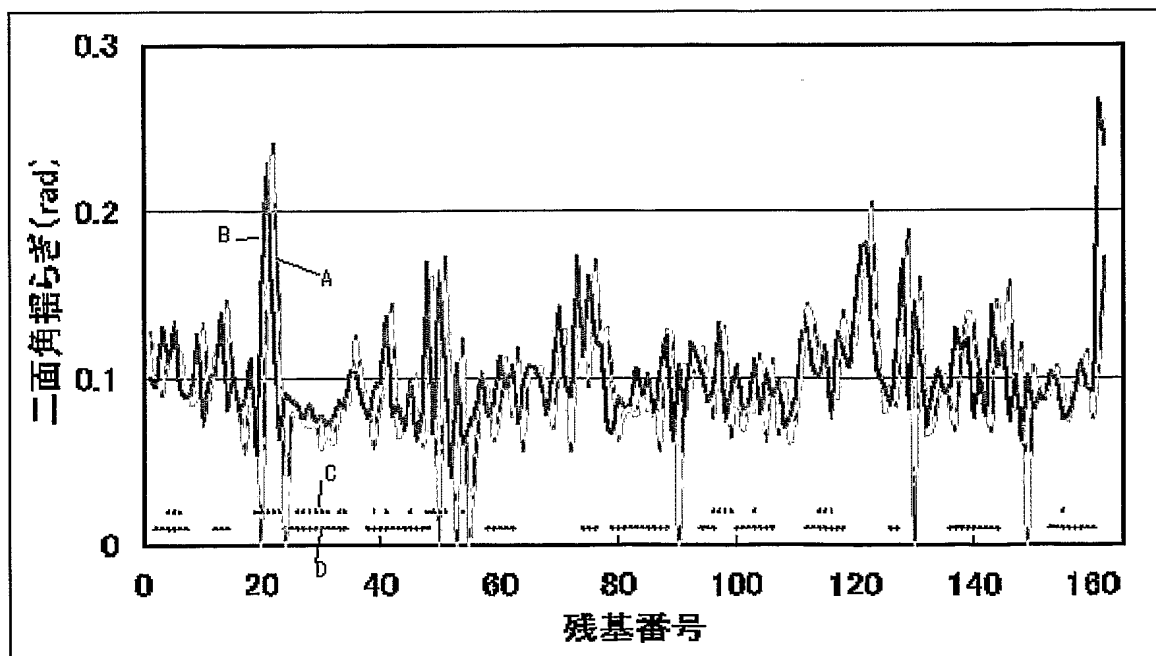
第9図



第10図



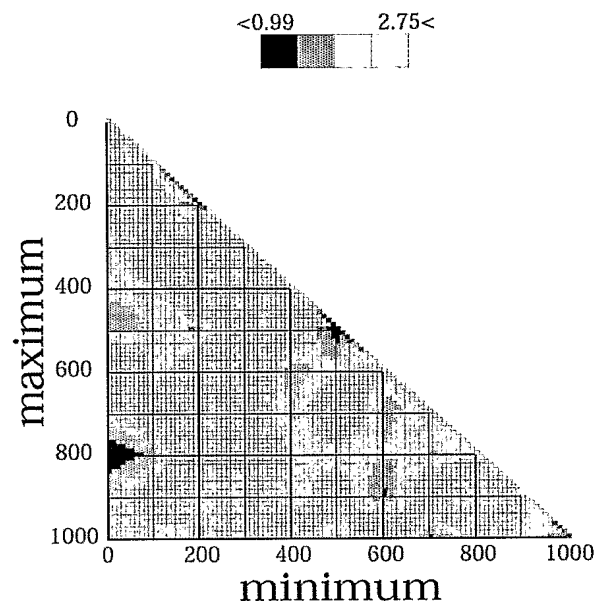
第11図



第12図

順位	最小値	最大値	α (%)	β (Å)	クラスター数	スコア
1	0	800	70	0.4	57	0.9054
2	0	800	70	0.1	62	0.9097
3	0	800	70	0.2	62	0.9097
4	0	800	70	0.3	62	0.9097
5	0	800	80	0.1	81	0.9102
6	0	800	80	0.2	81	0.9102
7	0	800	70	0.5	52	0.9103
8	0	800	80	0.4	73	0.9106
9	0	800	80	0.3	80	0.9116
10	0	800	80	0.5	67	0.9151
11	0	800	70	0.6	46	0.9156
12	0	800	90	0.5	240	0.9183
13	0	800	90	0.6	174	0.9194
14	0	800	60	0.6	13	0.9211
15	0	800	90	0.4	297	0.9225
16	0	800	80	0.6	58	0.9261
17	0	800	90	0.1	425	0.9286
18	0	800	90	0.2	425	0.9286
19	0	800	90	0.3	420	0.9296
20	0	800	60	0.1	16	0.9354
21	0	800	60	0.2	16	0.9354
22	0	800	60	0.3	16	0.9354
23	0	800	60	0.4	16	0.9354
24	0	800	60	0.5	15	0.9451
25	600	900	60	0.1	28	0.9469
26	600	900	60	0.2	28	0.9469
27	600	900	60	0.3	28	0.9469
28	600	900	60	0.4	28	0.9469
29	600	900	60	0.5	27	0.9518
30	600	900	60	0.6	27	0.9518

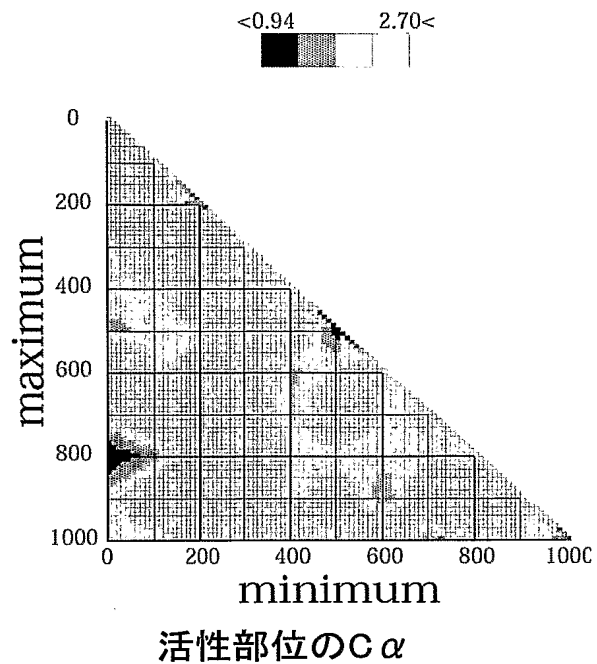
第13図



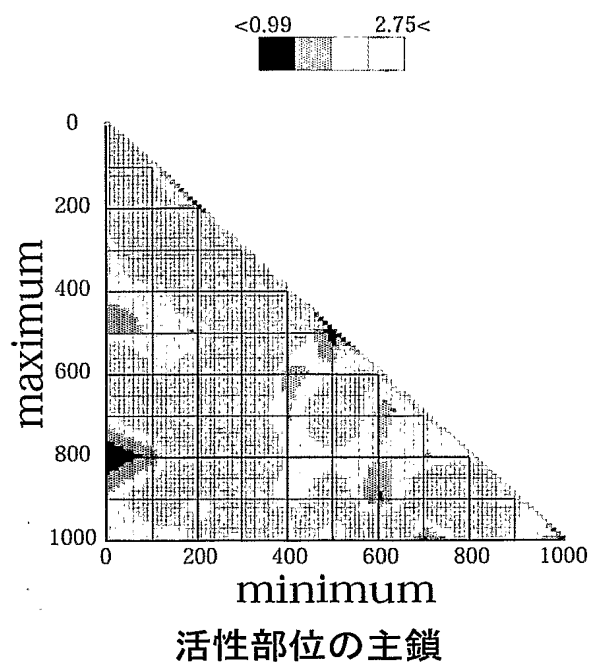
第14図

MD拘束の最小値	0.00
MD拘束の最大値	800.00
クラスタリングの定数 α (%)	80.00
クラスタリングの定数 β (Å)	0.40

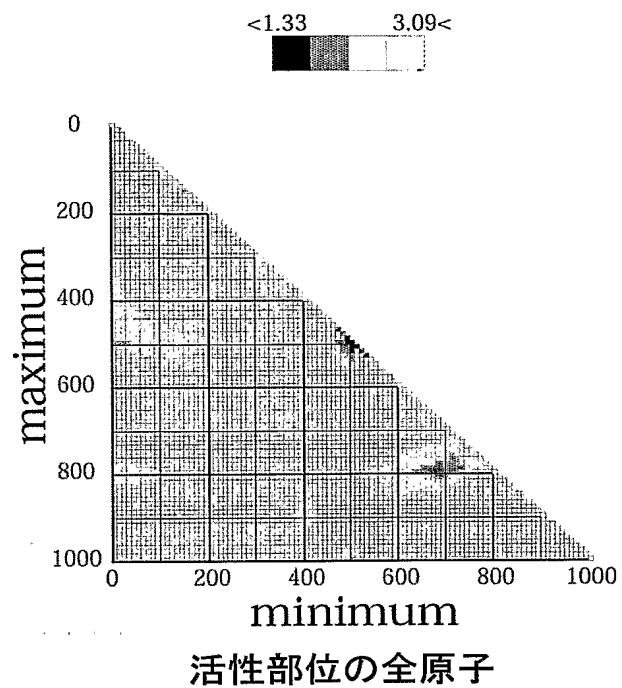
第15図



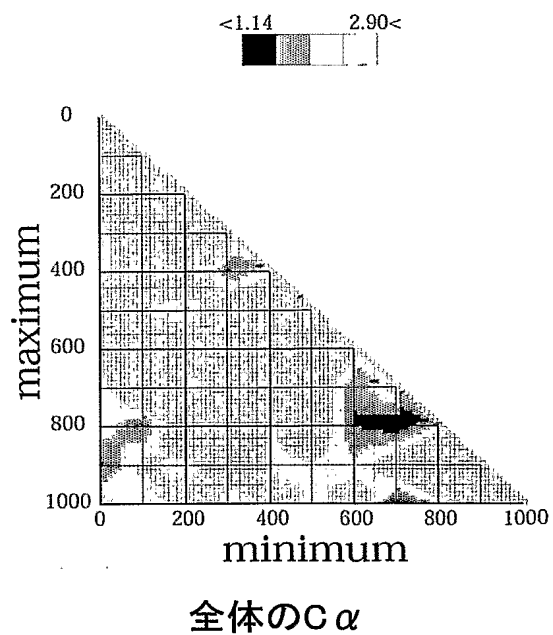
第16図



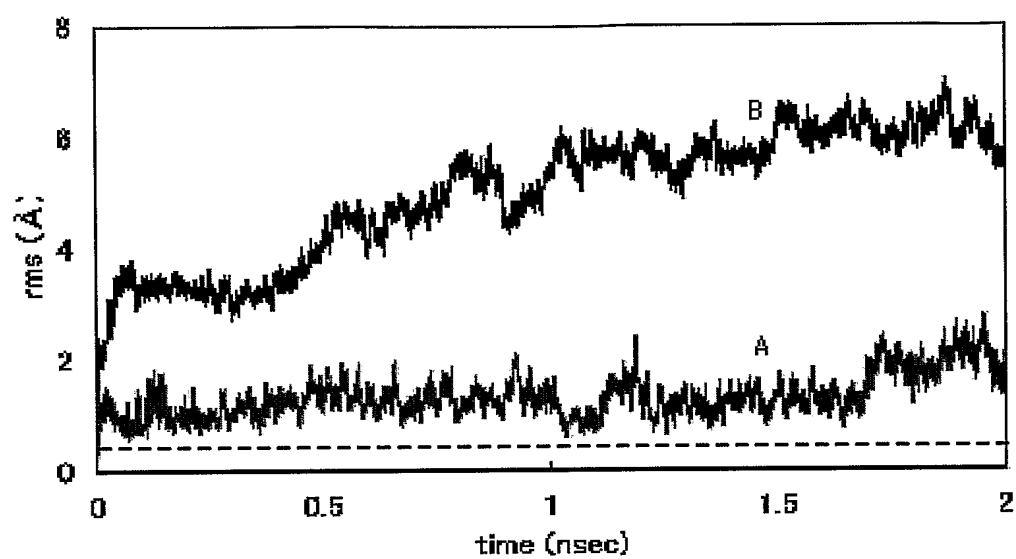
第17図



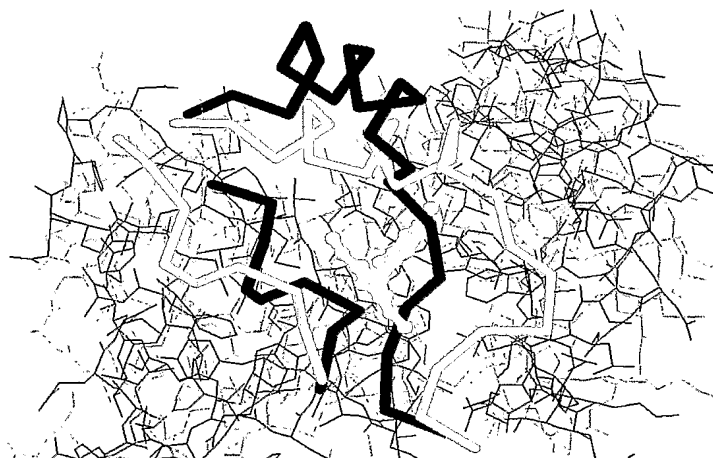
第18図



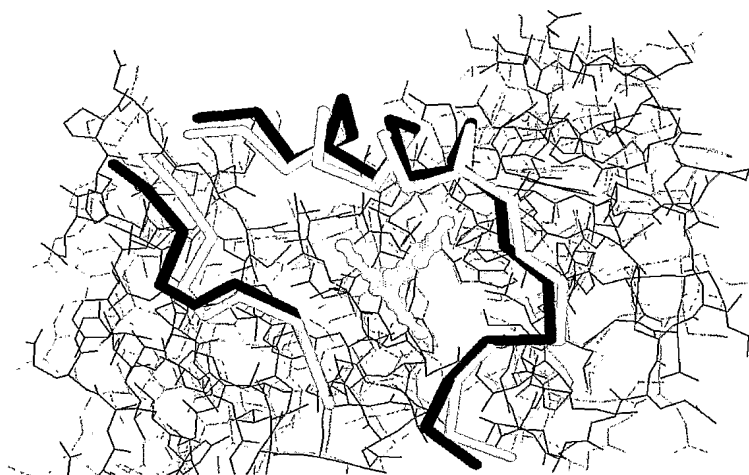
第19図



第20図



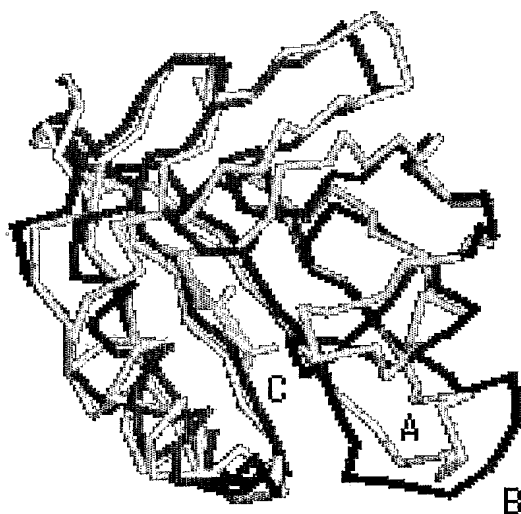
第21図



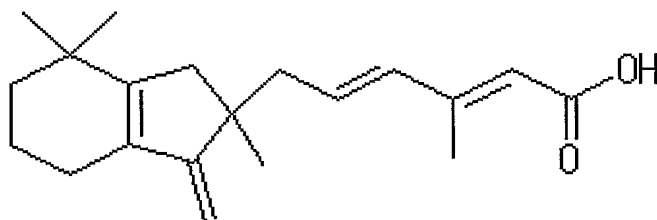
第22図

>1CBQ
PNFSGNWKIIRSENFEELLKVLGVNVMRLRKIAVAAASKPAVEIKQEGDTFYIKTSTTVRTTEINFKVGEFFEEQTVDGRP
CKSLVKWESENKMOVCEQKLLKGEGPKTSWTRELTNDGELILMTADDVVCTRVVYRE
>1ICM
-AFDGTWKVDRNENYEKFMKMGINVVKRKLG-AHDNLKLTTQEGNKFTVKESNFRNIDVVFELGVDFAYSLADGTE
L-TGTWTMEGNKLVGKFKRV-DNGKELJAVREIS-GNELIQTYTYEGVEAKRIFKKE

第23図



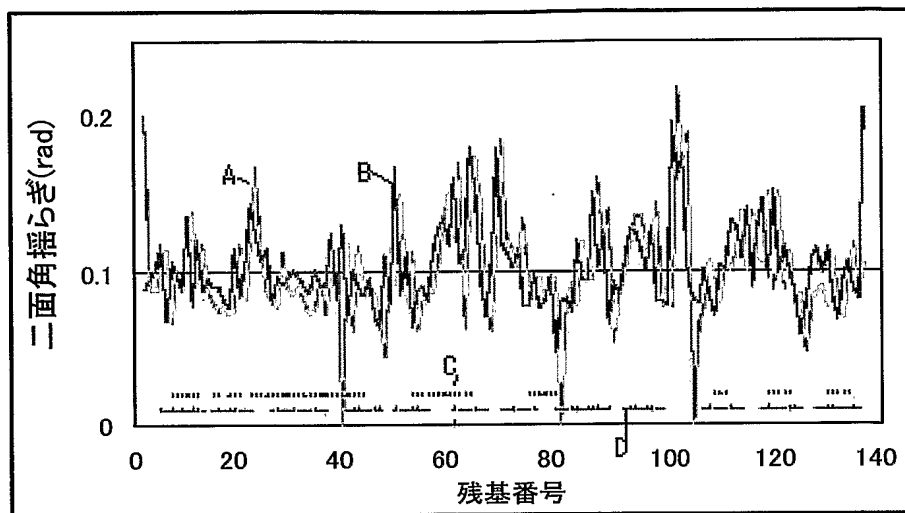
第24図



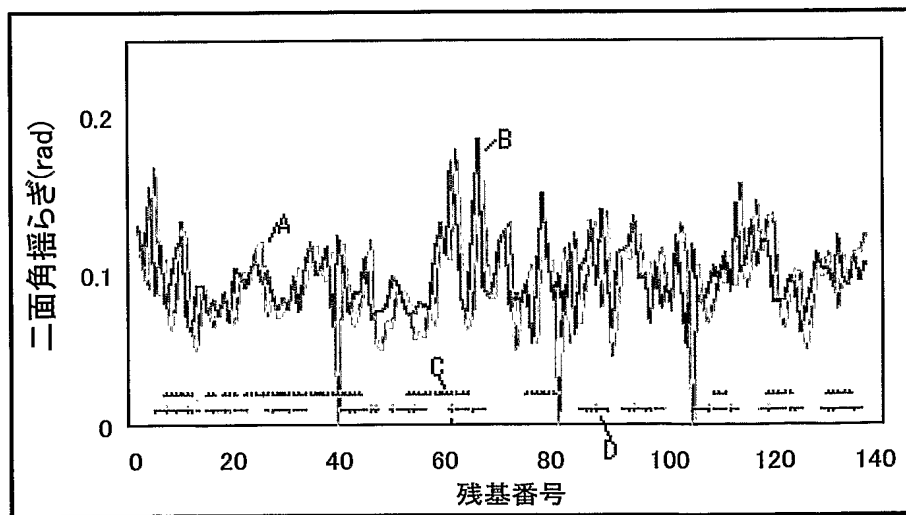
第25図

X線構造	1CBQ
参照タンパク質	1ICM
ホモロジー(%)	32.1
残基数	136
活性部位の主鎖(Å)	2.2487
活性部位の側鎖(Å)	3.2446
活性部位の全原子(Å)	2.7728
全体の主鎖(Å)	2.2075
全体の側鎖(Å)	3.7881
全体の全原子(Å)	3.0959

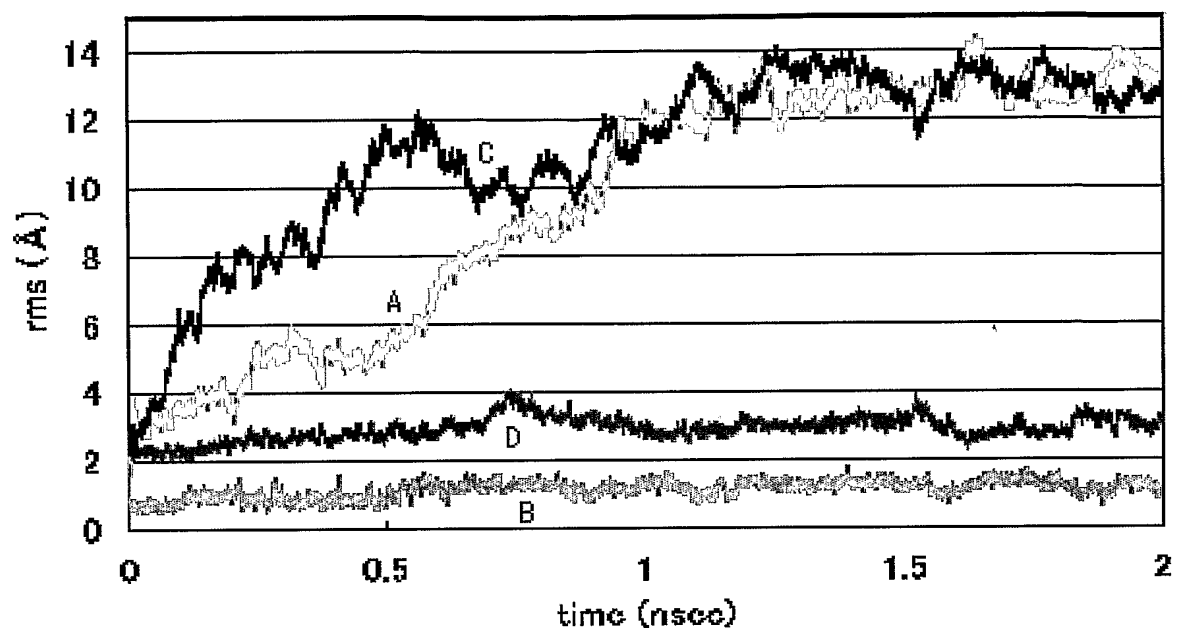
第26図



第27図



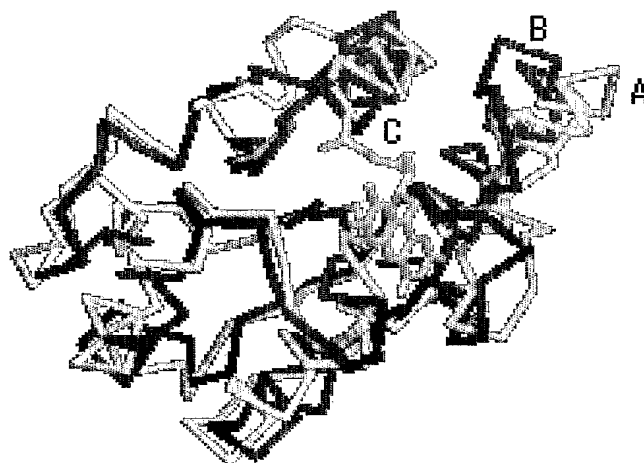
第28図



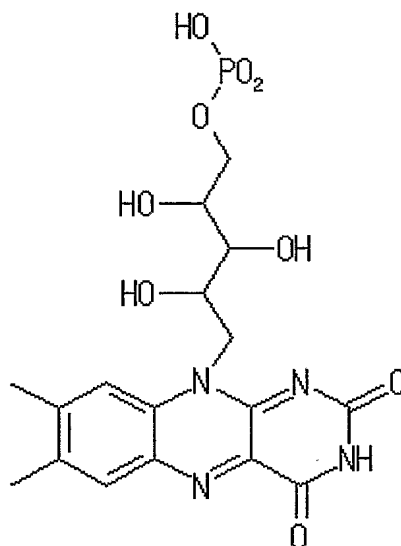
第29図

>1J9G
AKALIVYGSTTGNTTEYTAETIARELADAGYEVDSDRAASVEAGGLFEGFDLVLLGOSTWGD--DCIELQDDFIPLFDSLEE
TGAQGRKVAACFGCGDS--SYEYFCGAVDAIEEKLKNLGAETVQDG-----LRIDGDPRAARDDI
VGWAHDVVRGAI
>1AHN
AITGFFGSDTGNNTENIAKMIQKQLGKDVADVHDIKSSKE---DLEAYDILLGIPTWYYG---EAQCDWDDFFPTLEE
IDFNGKLVALFGCGDQEDYAEYFCDALGTIRDIIEPRGATVGHWPTAGYHFEASKGLADDDHFFVGLAIDEDRQPELTAE
RVEKWKQISE

第30図



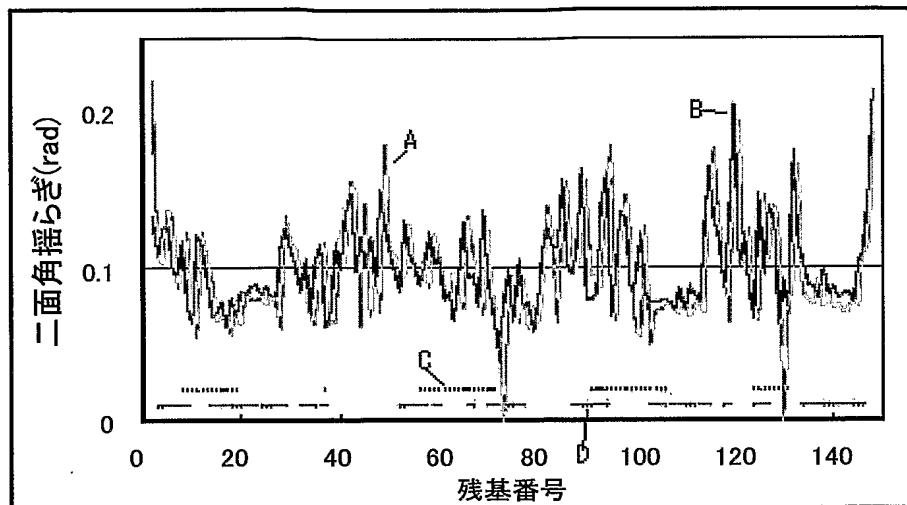
第31図



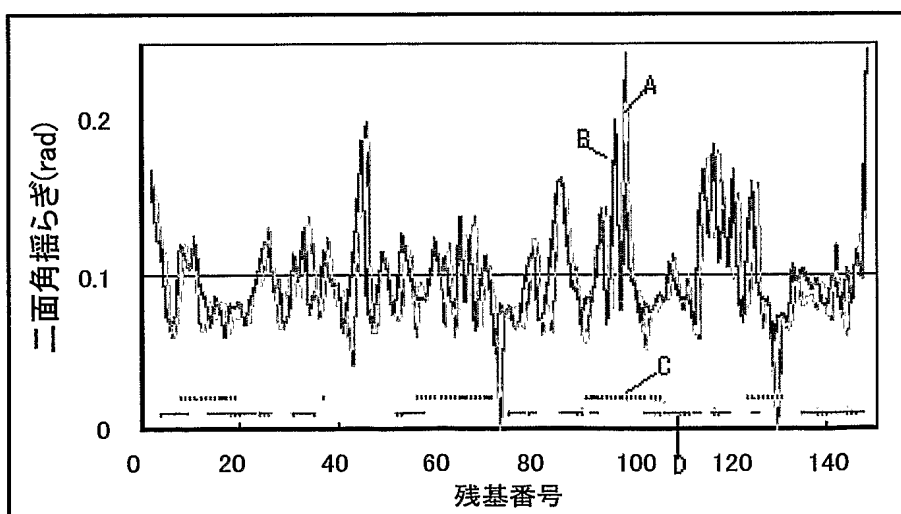
第32図

X線構造	1J9G
参照タンパク質	1AHN
ホモロジー(%)	29.2
残基数	147
活性部位の主鎖(Å)	2.3909
活性部位の側鎖(Å)	4.5774
活性部位の全原子(Å)	3.5753
全体の主鎖(Å)	3.1212
全体の側鎖(Å)	5.367
全体の全原子(Å)	4.315

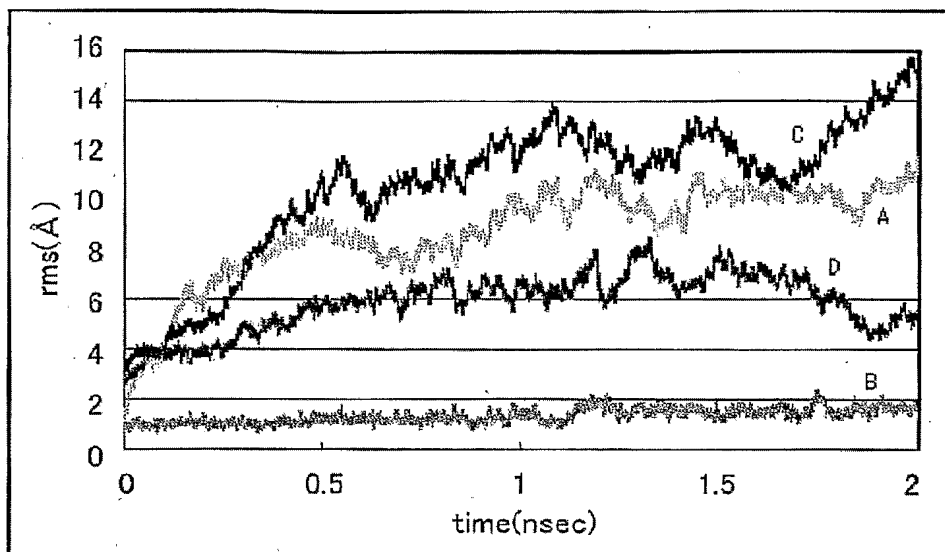
第33図



第34図



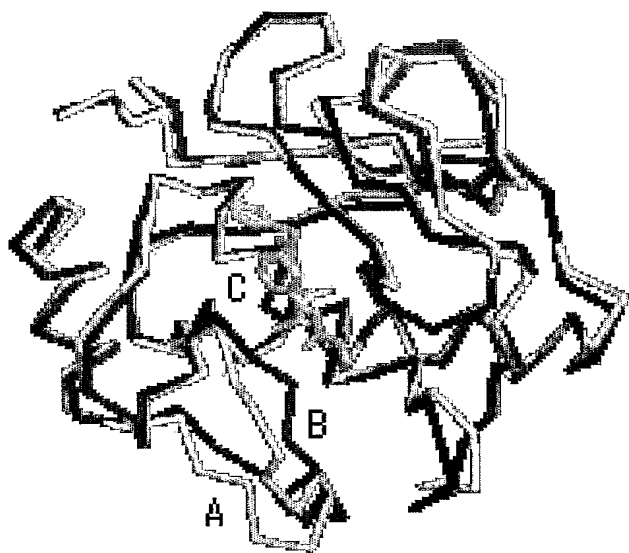
第35図



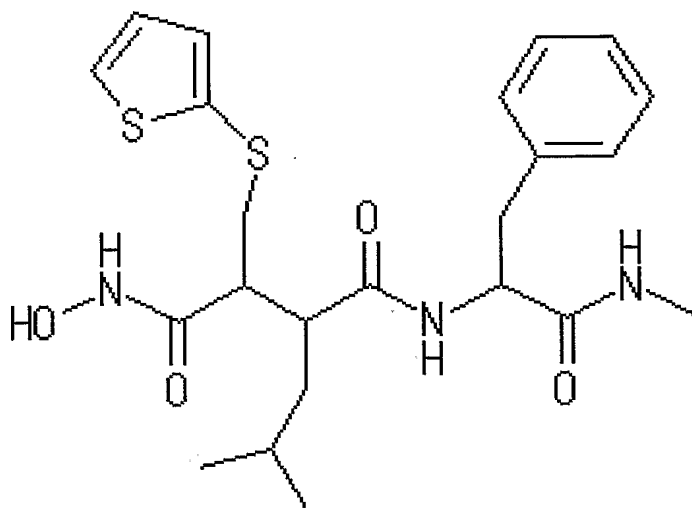
第36図

>1MMB
NPKWERTNLT YRINYTPQLSEAEVERAIKDAFELWSVASPLIFTRISQGEADINIAFYQRDHGDNSPFDGPNGLAHAF
QPGQGIGGDAHFDAEETWTNTSANYNLFLVAAHEFGHSLGLAHSSDPGALMYPNYA-FRETSNYSPLPQDDIDGIGAIYG
>1B3D_A
IPKWRKTHLT YRIVNYPDLPKDAVDSAVEKALKWEEVTPLTFSRLYEGEADIMISFAVREHGDFYPFDGPGNVLAHAY
APGPGINGDAHFDDDEQWTKDTTGTNLFLVAAHEIGHSLGLFHSANTEALMYPLYHSLTDLTRFRLSQDDINGIQSLYG

第37図



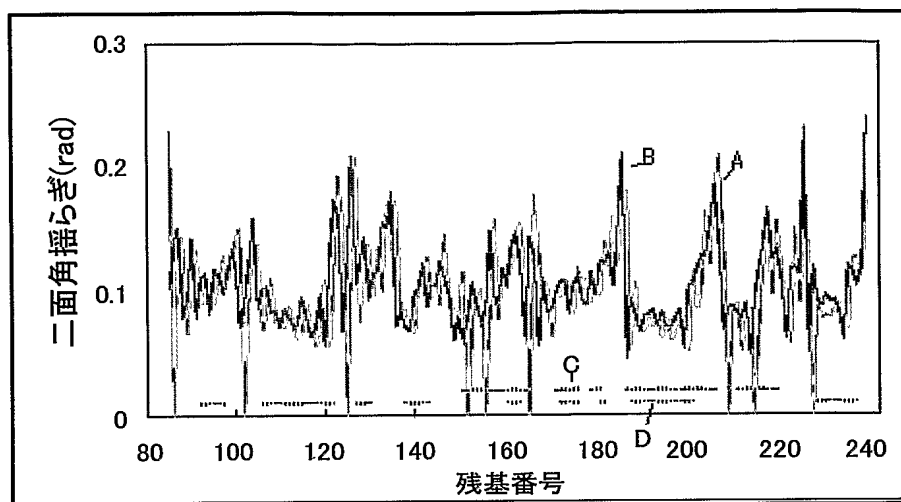
第38図



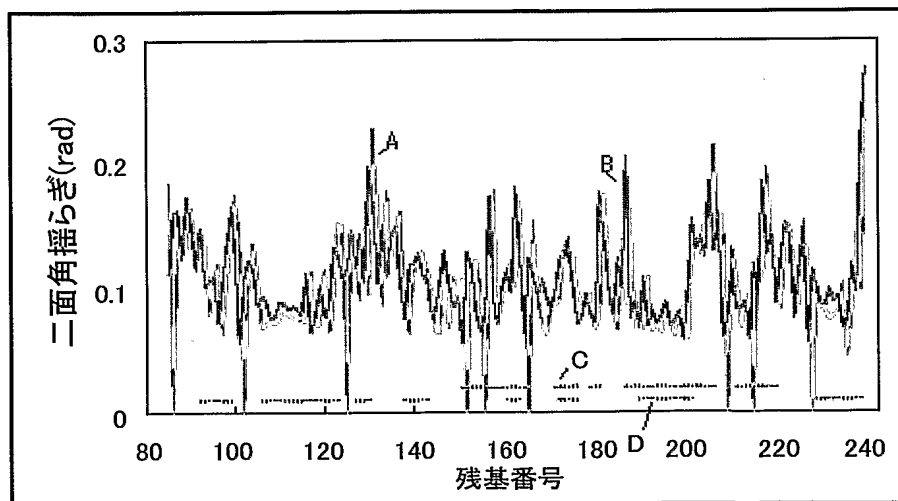
第39図

X線構造	1MMB
参照タンパク質	1B3D
ホモロジー(%)	55
残基数	158
活性部位の主鎖(Å)	0.9442
活性部位の側鎖(Å)	3.0756
活性部位の全原子(Å)	2.2417
全体の主鎖(Å)	1.1339
全体の側鎖(Å)	2.5715
全体の全原子(Å)	1.9808

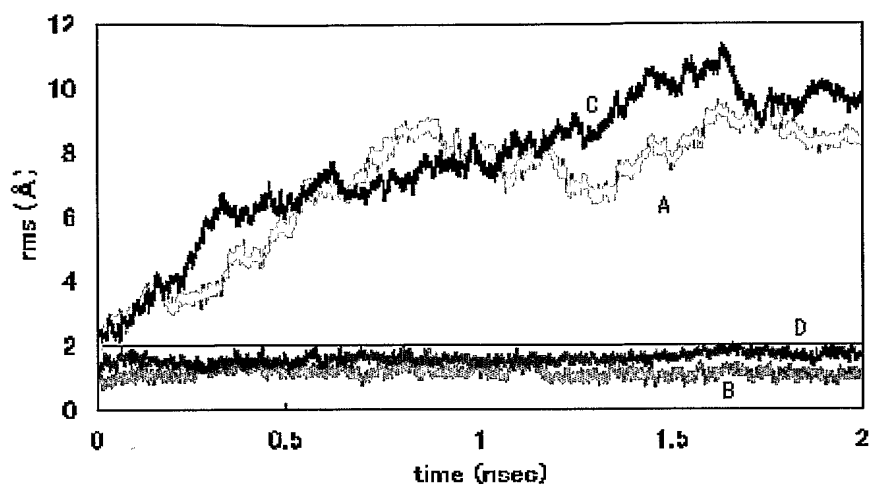
第40図



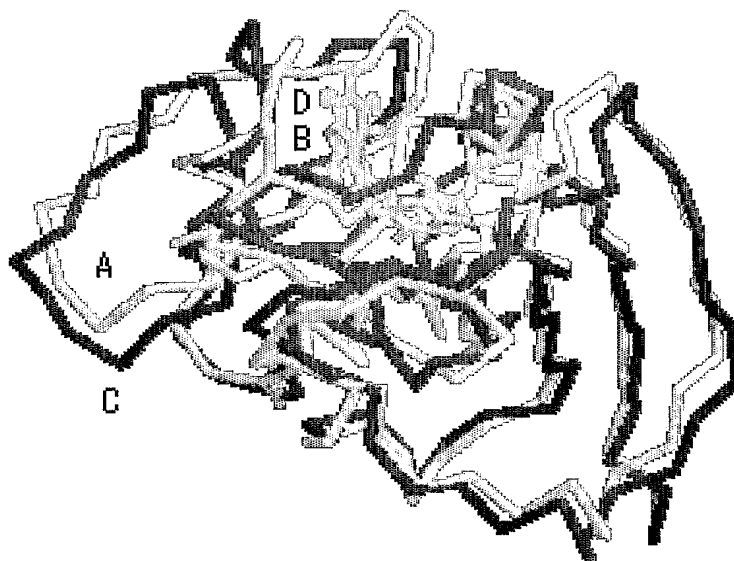
第41図



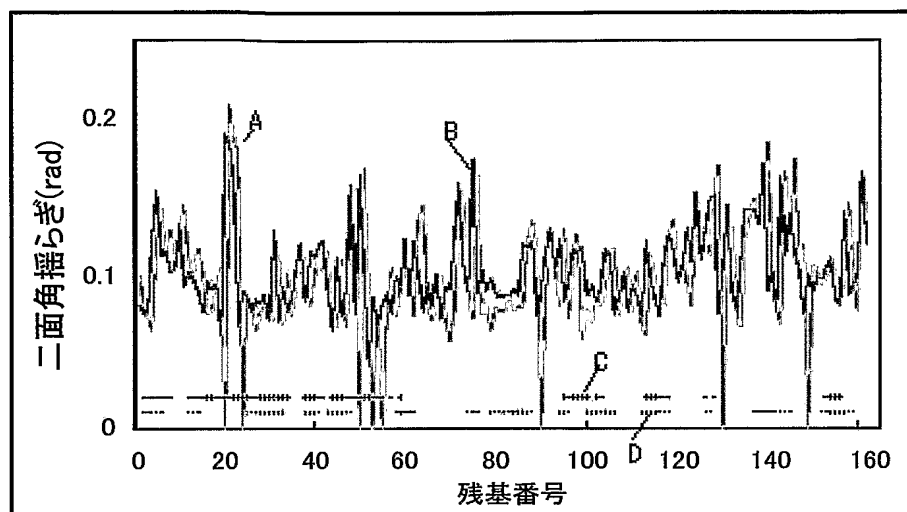
第42図



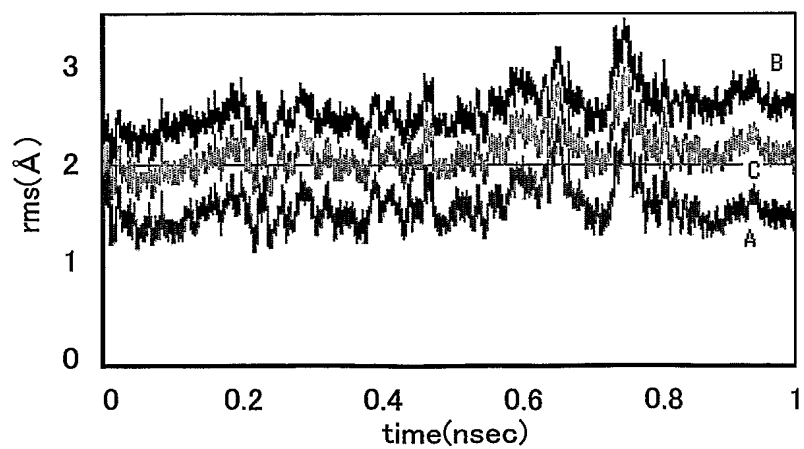
第43図



第44図



第45図



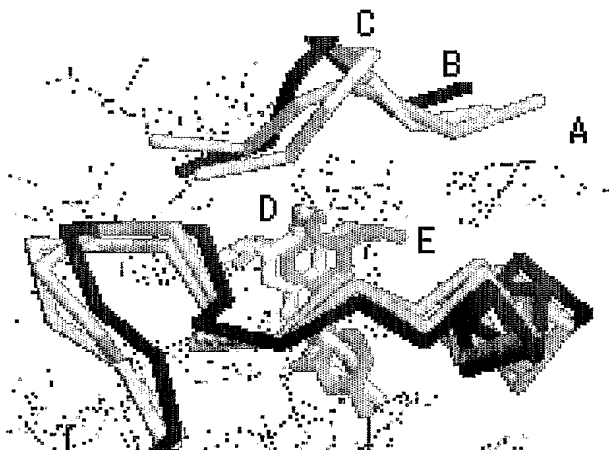
第46図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
LUE4 O	N.pl3	300	2.87
ASP26 OD1	N.ar	300	3.00
ASP26 OD2	N.pl3	300	3.00

第47図

区間(nsec)	間隔(fsec)	クラスター数	主鎖(Å)	全原子(Å)	リガンド(Å)
初期構造			1.5313	1.9190	
0~0.1	100	11	1.3531	1.8612	1.2734
0~1.0	100	204	1.2522	1.8116	0.9614
0~1.0	1000	26	1.2522	1.8116	0.8169

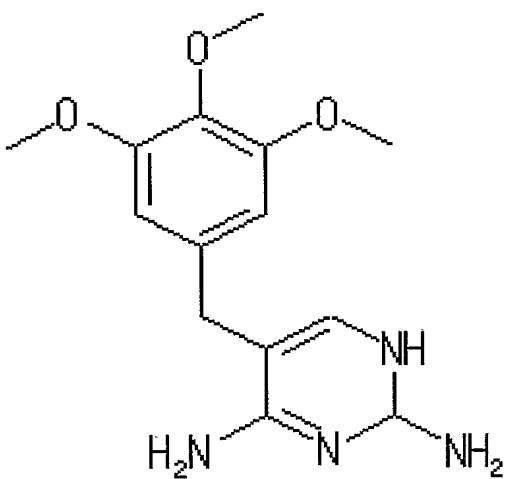
第48図



第49図



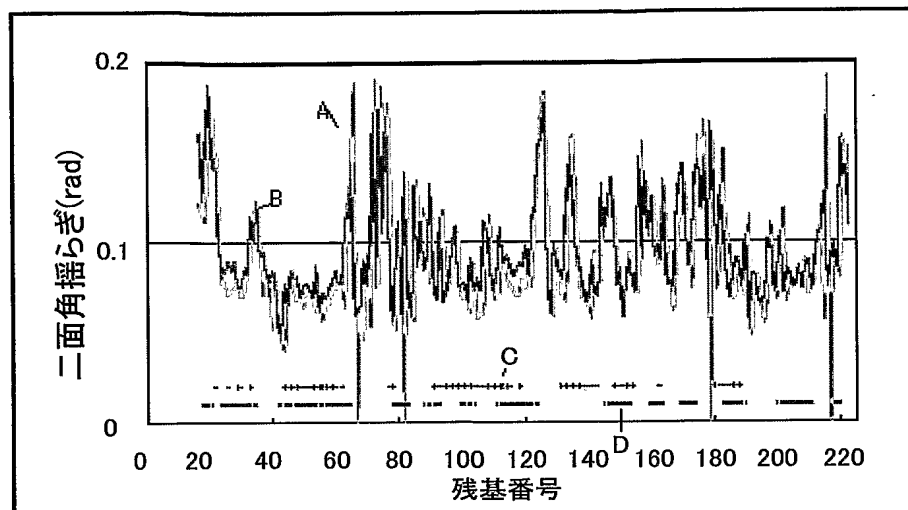
第50図



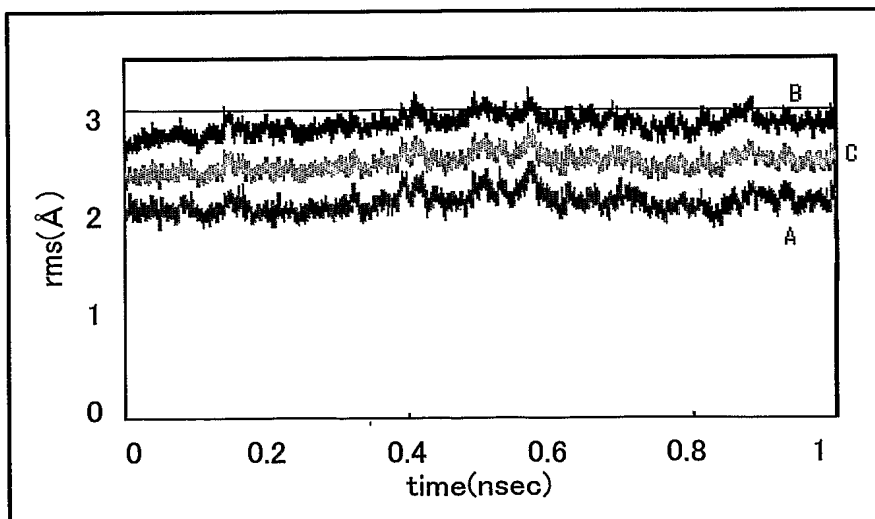
第51図



第52図



第53図



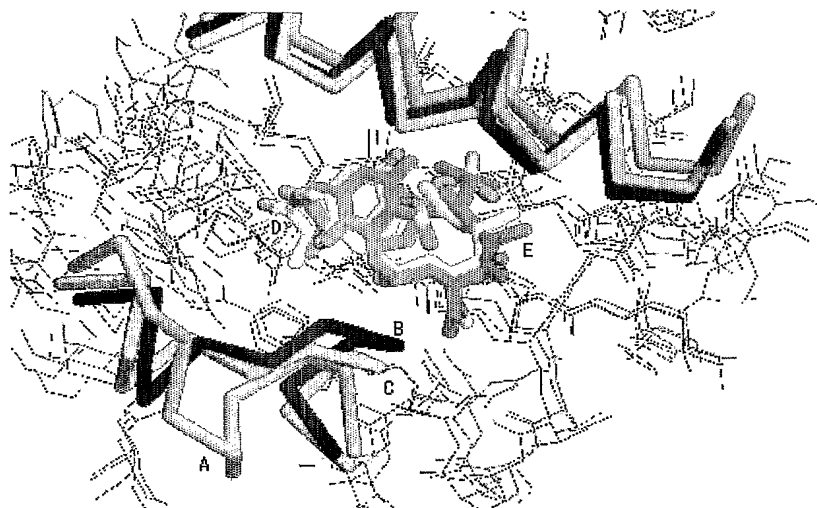
第54図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
LYS58 NZ	O.3	300	2.8
ASP93 OD2	N.am	300	2.8
PHE138 N	O.2	300	2.8

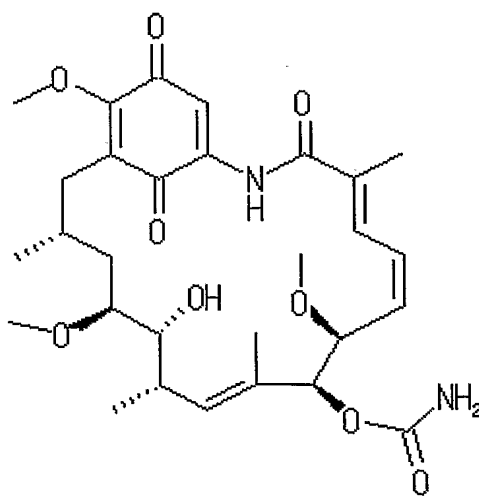
第55図

区間(nsec)	間隔(fsec)	クラスター数	主鎖(Å)	全原子(Å)	リガンド(Å)
初期構造			2.0144	2.2600	
0~0.1	100	6	1.8525	2.2601	1.2081
0~1.0	100	133	1.9139	2.3883	1.5932
0~1.0	1000	9	1.9764	2.8421	0.9667

第56図



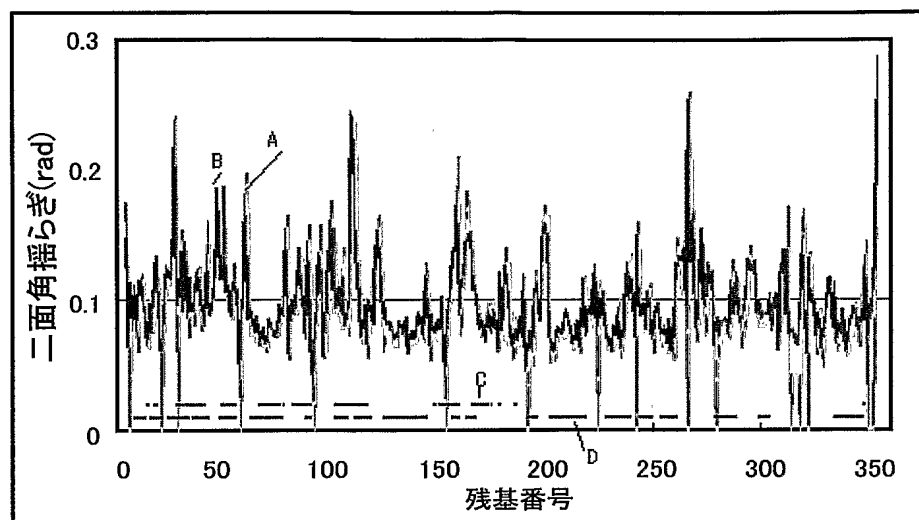
第57図



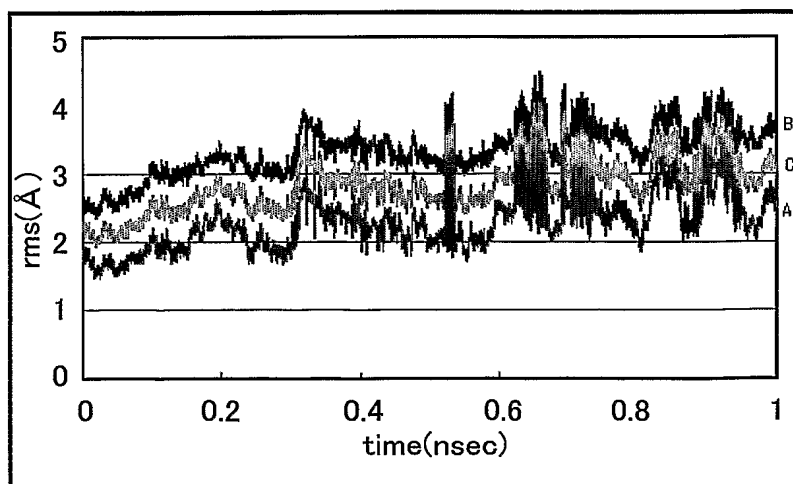
第58図



第59図



第60図



第61図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
LEU75 CD1	F	300	3.6
LEU75 CD2	F	300	3.6
MET109 N	Nar	300	2.7

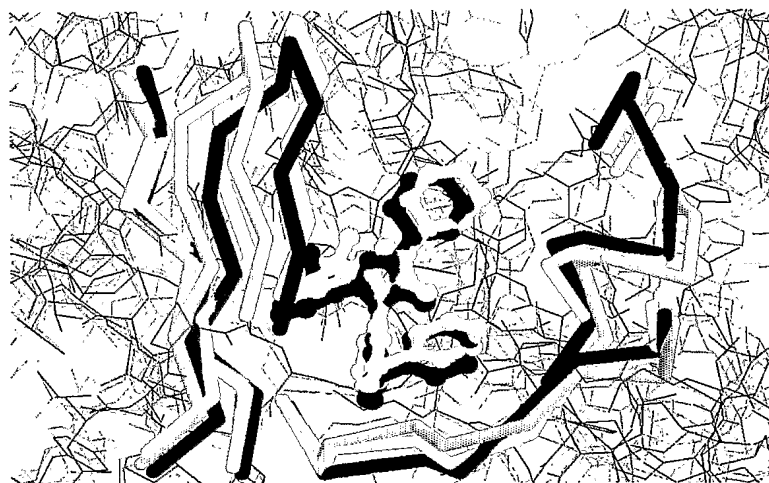
第62図

区間(nsec)	間隔(fsec)	クラスター数	主鎖(Å)	全原子(Å)	リガンド(Å)
初期構造			1.7972	2.1606	
0~0.1	100	5	1.6101	2.0766	1.6112
0~1.0	100	319	1.7236	2.2843	1.4550
0~1.0	1000	31	1.7236	2.2843	1.4571

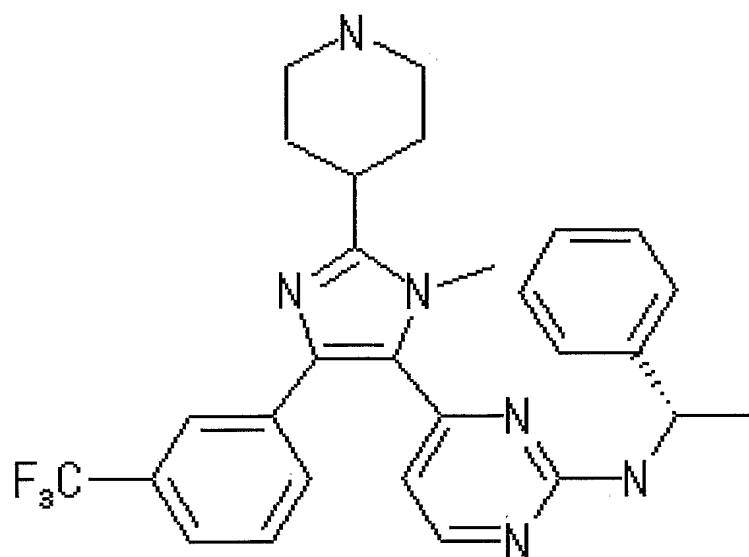
第63図



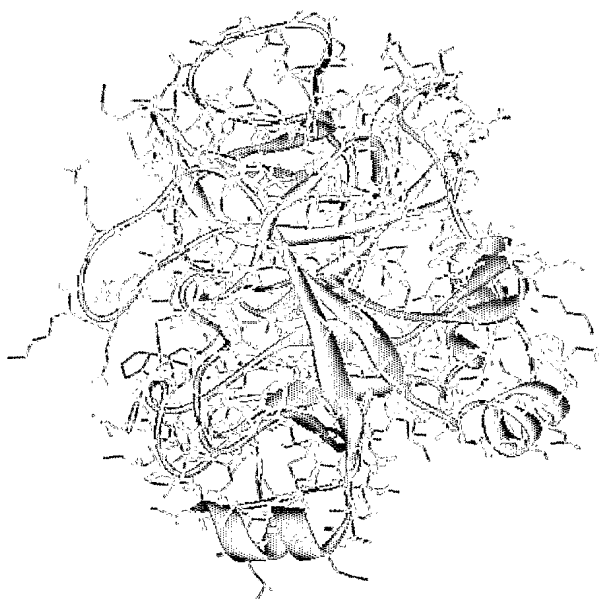
第64図



第65図



第66図



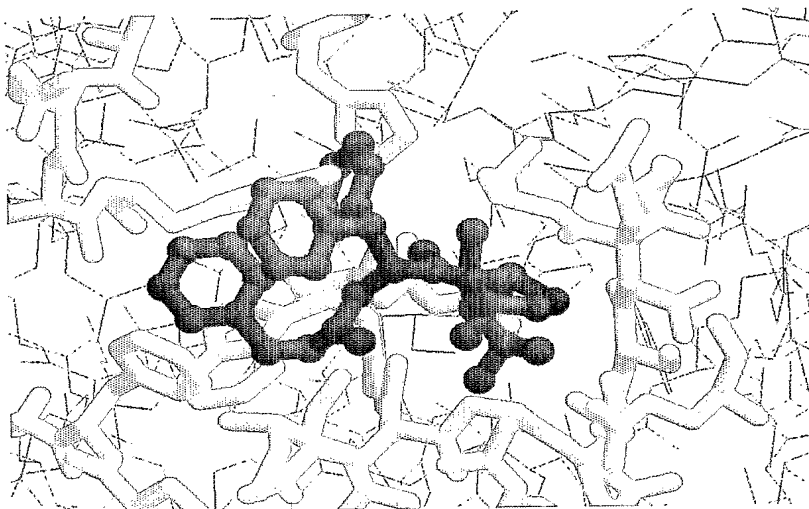
第67図

ランキング	相互作用エネルギー	PDBcode	ランキング	相互作用エネルギー	PDBcode
[001]	-3847.2147	4PGT	[002]	-3671.4754	1MP7
[003]	-3056.6135	1BMN	[004]	-2923.8680	1I32
[005]	-2872.4420	5FWG	[006]	-2608.5702	1LHF
[007]	-2528.6110	1BX6	[008]	-2439.5657	1B8Y
[009]	-2433.9052	1EZF	[010]	-2382.8539	5LDH
[011]	-2248.0139	1FVP	[012]	-2247.3089	1JJQ
[013]	-2133.5942	1IL2	[014]	-2128.4540	1BJI
[015]	-2125.1405	1DMT	[016]	-2103.1434	1K22
[017]	-2092.6654	1HY7	[018]	-2025.5091	966C
[019]	-2013.9064	1AIX	[020]	-1989.1635	1A4Q
[021]	-1946.4497	1VZE	[022]	-1932.9896	1KVO
[023]	-1928.3650	1D6V	[024]	-1901.6172	1C0A
[025]	-1890.2208	1DB5	[026]	-1867.0754	1GUH
[027]	-1855.6184	1QIN	[028]	-1817.4767	1M21
[029]	-1782.5387	1KCI	[030]	-1766.9010	1KZK
[031]	-1728.2876	6GSX	[032]	-1709.9359	2PRG
[033]	-1699.2351	1NPW	[034]	-1694.4086	2UPJ
[035]	-1661.4315	1AUJ	[036]	-1658.1970	1HFR
[037]	-1654.2430	1DMP	[038]	-1599.5870	1FOR
[039]	-1595.7907	2GSQ	[040]	-1569.9256	1QHC
[041]	-1530.3871	1AIM	[042]	-1481.1846	1EL3
[043]	-1473.7372	1QH5	[044]	-1453.3935	1LHC
[045]	-1411.1465	1HFC	[046]	-1389.8129	2FMB
[047]	-1372.1506	1GFW	[048]	-1352.8868	1EM6
[049]	-1329.5658	1AU0	[050]	-1306.5704	1M9B
[051]	-1287.3729	1EAS	[052]	-1265.8962	1LHE
[053]	-1248.8527	1C8T	[054]	-1244.2458	1MMQ
[055]	-1216.6454	1QIP	[056]	-1200.9810	207D
[057]	-1175.5120	1HWL	[058]	-1138.1881	4UPJ
[059]	-1112.7163	3GST	[060]	-1068.0641	1LEE
[061]	-1030.5972	1GA9	[062]	-1030.4960	1OD7
[063]	-1029.0345	1HOV	[064]	-1018.1686	1LF2
[065]	-1011.9100	1ODY	[066]	-976.1041	1CQQ
[067]	-948.0992	1G2K	[068]	-936.9058	2AIM
[069]	-934.4739	1NWL	[070]	-924.6255	6FIV
[071]	-902.7587	1YEI	[072]	-900.4131	1MXT
[073]	-894.5544	1YEF	[074]	-874.9274	1DZT
[075]	-857.5373	1QF0	[076]	-851.1669	1EGV
[077]	-844.2406	1F29	[078]	-824.5393	1KV2
[079]	-820.4913	456C	[080]	-775.9659	1K1M
[081]	-766.8359	1JR4	[082]	-763.2825	2KOE
[083]	-739.3676	1KN4	[084]	-733.8593	1RT2
[085]	-728.8765	1HPV	[086]	-718.5795	2BBQ
[087]	-705.3978	1MS6	[088]	-695.0241	1IF7
[089]	-689.7998	1JIL	[090]	-684.7289	1A8J
[091]	-676.3861	1FL3	[092]	-628.8081	1CIZ
[093]	-619.2121	1DIF	[094]	-604.7057	2BPX
[095]	-598.4143	1IF9	[096]	-564.5807	1K0C
[097]	-561.6472	1KN2	[098]	-541.1021	1HBV
[099]	-507.6808	1DB4	[100]	-496.0550	1K1J

太字: 1 AIX中に含まれているリガンド

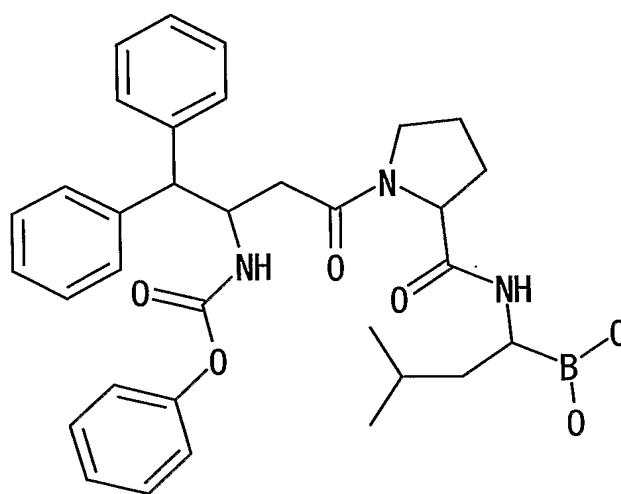
斜線: セリンプロテアーゼ

第68図



ランキング19

第69図



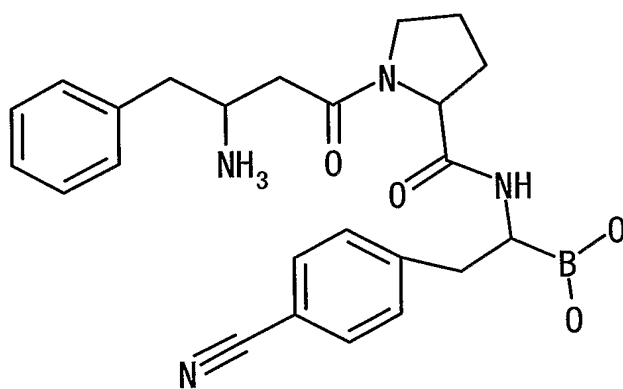
1AIX中のリガンド

第70図



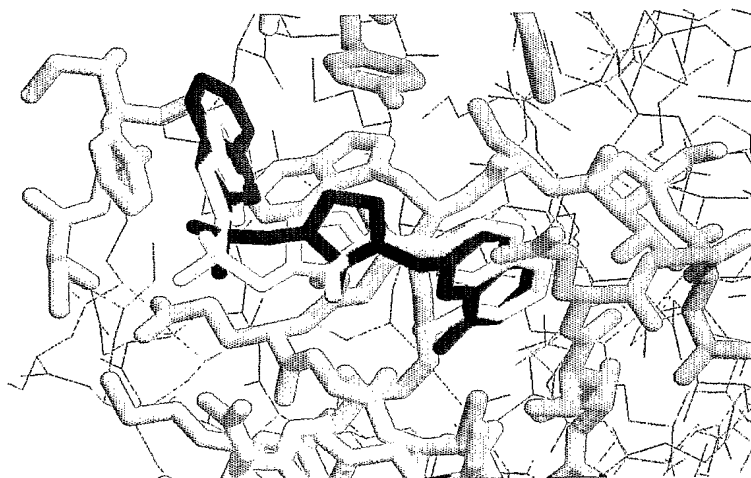
ランキング35

第71図



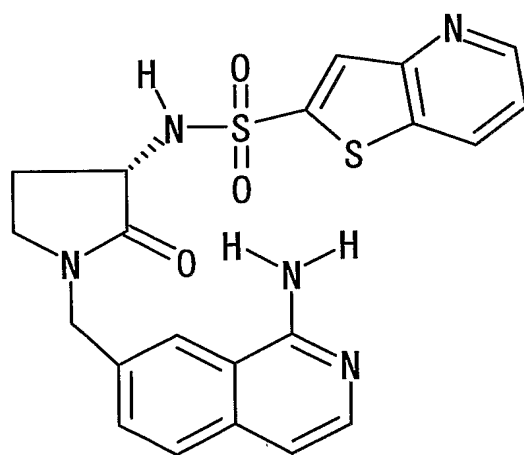
1AUJ中のリガンド

第72図



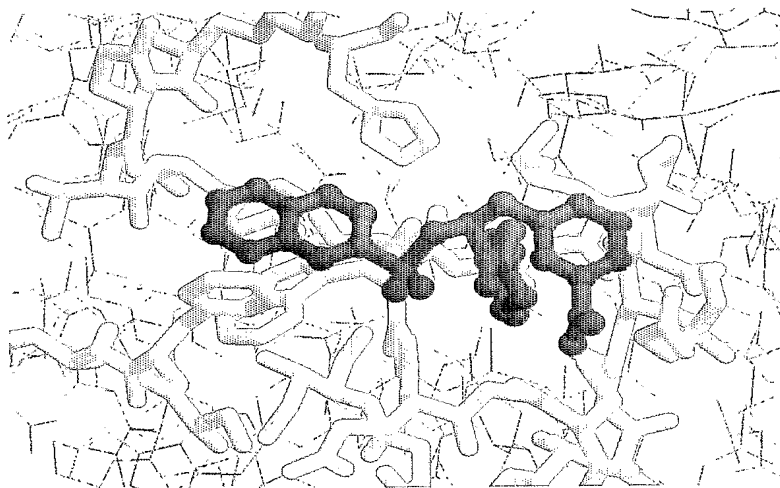
ランキング38

第73図



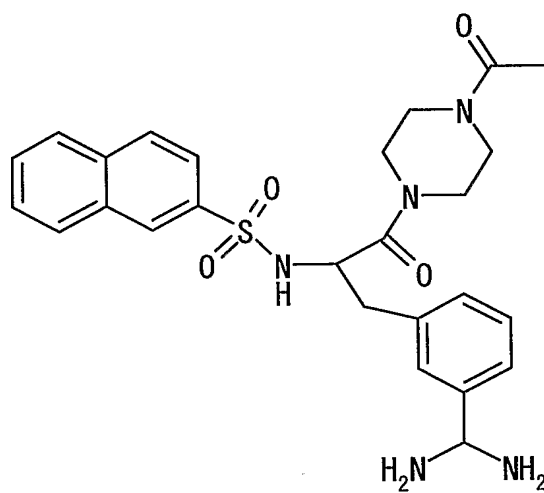
1FOR中のリガンド

第74図



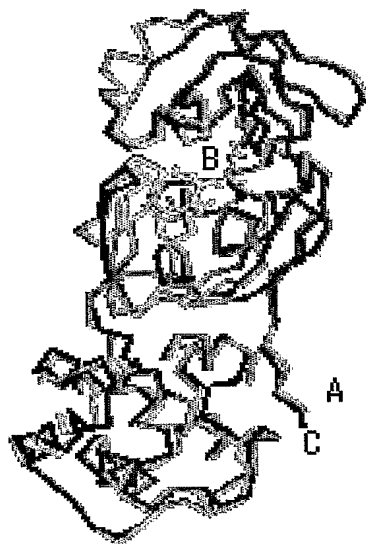
ランキング80

第75図

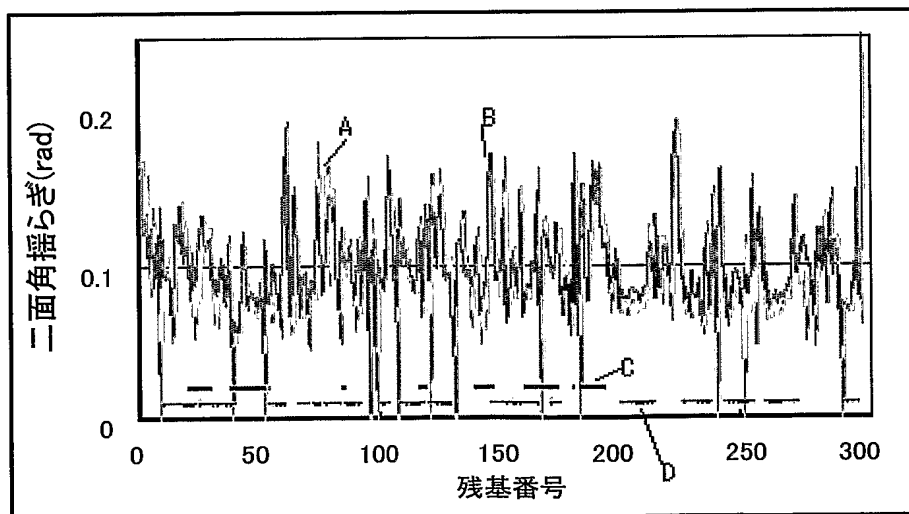


1KIM中のリガンド

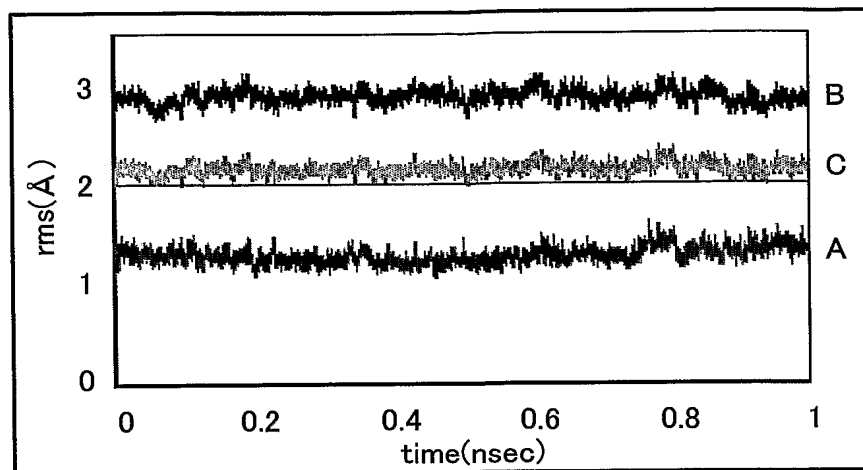
第76図



第77図



第78図



第79図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
CYS145 N	O.co2	100	2.70
MET165 CG	C.3	100	4.00
GLU166 N	O.2	100	2.70
THR190 N	O.3	100	2.70

第80図

順位	エネルギー	PDB code	備考
1	-1089.2153	1QF4	ligase
2	-990.9917	1KZL	transferase
3	-906.5003	1C0A	ligase/RNA
4	-889.1661	1KGQ	transferase
5	-869.3531	1I95	ribosome
6	-860.2331	1JR4	transferase
7	-858.0005	1A2N	transferase
8	-832.0515	1NKK	hydrolase
9	-788.3545	1JIL	ligase
10	-757.2852	1EJB	transferase
11	-697.9477	1DMT	hydrolase
12	-645.0269	1PAU	complex (protease/inhibitor)
13	-633.1260	1F74	lyase
14	-628.9678	1KYU	endocytosis/exocytosis
15	-616.4458	1NRS	serine proteinase/receptor
16	-608.4169	9LYZ	hydrolase (o-glycosyl)
17	-600.2775	1EIO	lipid-binding protein
18	-593.7082	1F7B	lyase
19	-585.7663	1LMW	complex (serine protease/inhibitor)
20	-584.0059	1R1R	oxidoreductase
21	-580.1563	1IL2	ligase/RNA
22	-573.0481	1BLL	hydrolase(alpha-aminoacylpeptide)
23	-572.6763	1E1F	glycoside hydrolase
24	-540.1965	1LKL	complex (tyrosine kinase/peptide)
25	-524.2817	1UK4	hydrolase
26	-518.3528	1LCB	transferase (methyltransferase)
27	-506.8123	1PGN	oxidoreductase (choh(d)-nadp+(a))
28	-493.5477	1I5Q	hydrolase
29	-486.8954	1KYD	endocytosis/exocytosis
30	-481.9659	1NRR	serine proteinase/receptor

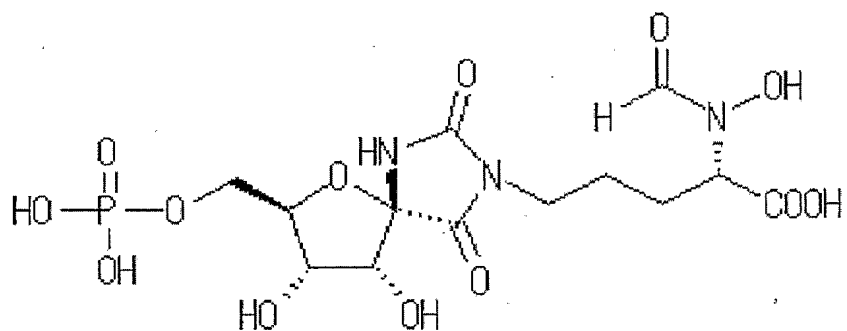
第81図



第82図



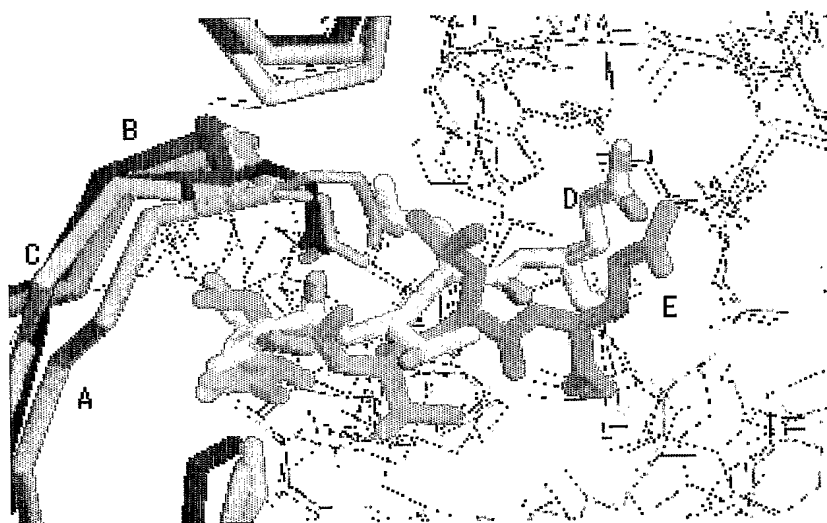
第83図



第84図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
CYS145 N	O.co2	100	2.70
GLUの166 N	O.2	100	2.70
THR190 N	O.3	100	2.70

第85図



第86図

順位	エネルギー	PDB code	備考
1	-1263.8870	1EAD	dihydrolipoamide acetyltransferase
2	-1260.8689	1F6M	oxidoreductase
3	-1147.1739	1JR4	transferase
4	-1141.9917	1QF5	ligase
5	-1104.9447	1JAY	structural genomics
6	-1019.3584	1KZL	transferase
7	-996.5865	1QF4	ligase
8	-988.6588	1JIJ	ligase
9	-981.8594	8ICO	complex (nucleotidyltransferase/dna)
10	-953.0986	1LO9	hydrolase
11	-949.1903	1JTU	transferase
12	-922.4795	1JKX	transferase
13	-918.4892	1JIL	ligase
14	-916.9950	1I95	ribosome
15	-908.4880	1AL6	lyase
16	-893.5862	1LKL	complex (tyrosine kinase/peptide)
17	-892.3713	1N37	deoxyribonucleic acid
18	-887.9721	1LCB	transferase (methyltransferase)
19	-866.9600	1O9F	protein binding
20	-826.4893	1LO7	hydrolase
21	-792.0254	4UAG	ligase
22	-776.9998	1EJB	transferase
23	-772.2400	1BFZ	n-terminal product peptide
24	-769.6844	1F9E	apoptosis
25	-762.5275	1TLP	hydrolase (metalloproteinase)
26	-759.8312	1QIN	lyase
27	-758.2140	1KO6	transferase
28	-757.5526	1C0A	ligase/RNA
29	-755.7987	1QD1	transferase
30	-755.1049	1LO8	hydrolase
49	-639.1858	1UK4	hydrolase

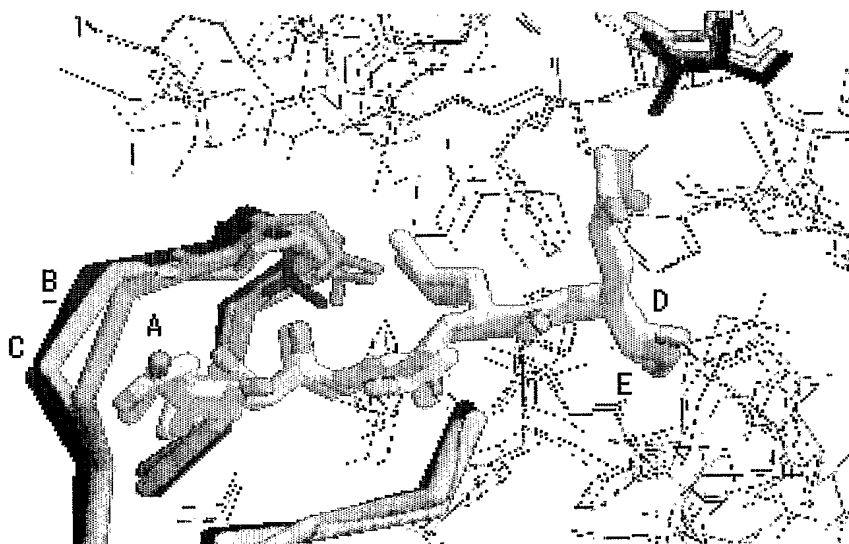
第87図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
THR25 OG1	N.am	100	3.80
CYS145 N	O.co2	100	2.70
MET165 OG	C.3	100	4.00
GLU168 N	O.2	100	2.70
THR190 N	O.3	100	2.70

第88図

順位	エネルギー	PDB code	備考
1	-364.6548	1I95	ribosome
2	-299.0166	1UK4	hydrolase
3	-109.6867	1BXX	endocytosis/exocytosis
4	-93.0540	1KZL	transferase
5	-72.9399	1NKK	hydrolase
6	-10.7565	1F8H	endocytosis/exocytosis
7	-4.2756	1QTN	apoptosis
8	162.1557	1KGQ	transferase
9	163.2075	1O9F	protein binding
10	331.8725	1CGL	metalloprotease
11	370.5027	2BBQ	transferase(methyltransferase)
12	397.8488	4DMR	oxidoreductase
13	550.2598	1HPG	hydrolase (serine protease)
14	716.6561	1LOC	lectin
15	839.7398	1DMT	hydrolase
16	848.7090	1KAP	zinc metalloprotease
17	850.2630	1JG3	transferase
18	883.4400	1BC5	complex (methyltransferase/peptide)
19	905.9695	1FCH	signaling protein
20	913.9769	1CF8	catalytic antibody
21	1088.2428	1NWE	hydrolase
22	1089.3496	1KO6	transferase
23	1116.9042	1F74	lyase
24	1131.4783	1ING	hydrolase (o-glycosyl)
25	1132.3648	1I31	endocytosis/exocytosis
26	1148.9063	1IAU	hydrolase
27	1156.0335	1B48	transferase
28	1160.3512	1PTT	complex (hydrolase/peptide)
29	1176.7814	1MC5	oxidoreductase
30	1197.3565	1F9E	apoptosis

第89図



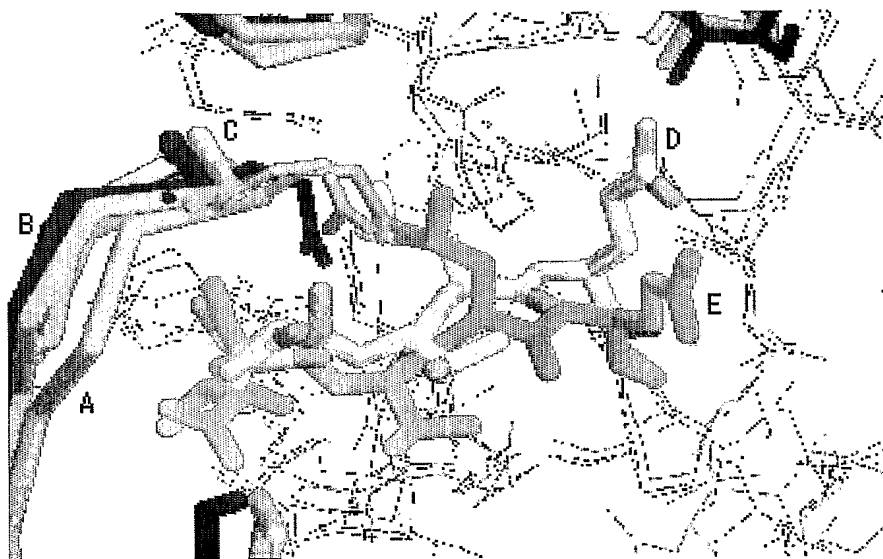
第90図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
CYS145 N	アクセプター	100	2.70
MET165 CG	炭素	100	4.00
GLU166 N	アクセプター	100	2.70
THR190 N	アクセプター	100	2.70

第91図

順位	エネルギー	PDB code	備考
1	-2095.8588	1JJQ	hormone/growth factor
2	-2011.3626	2BVW	hydrolase
3	-1670.8384	1DOG	hydrolase
4	-1336.7960	1LWJ	transferase
5	-1320.0704	1KEU	lyase
6	-1230.0604	1GAH	hydrolase
7	-1214.9459	1I7E	signaling protein
8	-1195.8653	1C39	signaling protein
9	-1191.3777	1BB5	hydrolase
10	-1189.0253	2FHI	nucleotide-binding protein
11	-1147.9761	1GO6	glycopeptide antibiotics
12	-1103.6272	1M4D	transferase
13	-1095.3050	1QHC	hydrolase
14	-1088.7299	1M2N	gene regulation
15	-1078.3684	1QGL	lectin (agglutinin)
16	-1056.4078	4ENG	glycosyl hydrolase
17	-1033.0227	1LON	ligase
18	-1031.2555	1MWL	ribonucleic acid
19	-1027.4239	1QPK	hydrolase
20	-1014.9817	1UDB	isomerase
21	-1005.1689	1GQC	transferase
22	-976.9293	1H6H	px domain
23	-975.2827	1LSP	hydrolase (o-glycosyl)
24	-973.5218	1FF1	signaling protein
25	-963.4098	3UAG	ligase
26	-937.2165	1IBG	immunoglobulin
27	-933.6818	1DRV	oxidoreductase
28	-918.6947	2MBR	oxidoreductase
29	-917.1703	1NAB	deoxyribonucleic acid
30	-897.3026	1SLY	glycosyltransferase
774	331.9928	1UK4	hydrolase

第92図



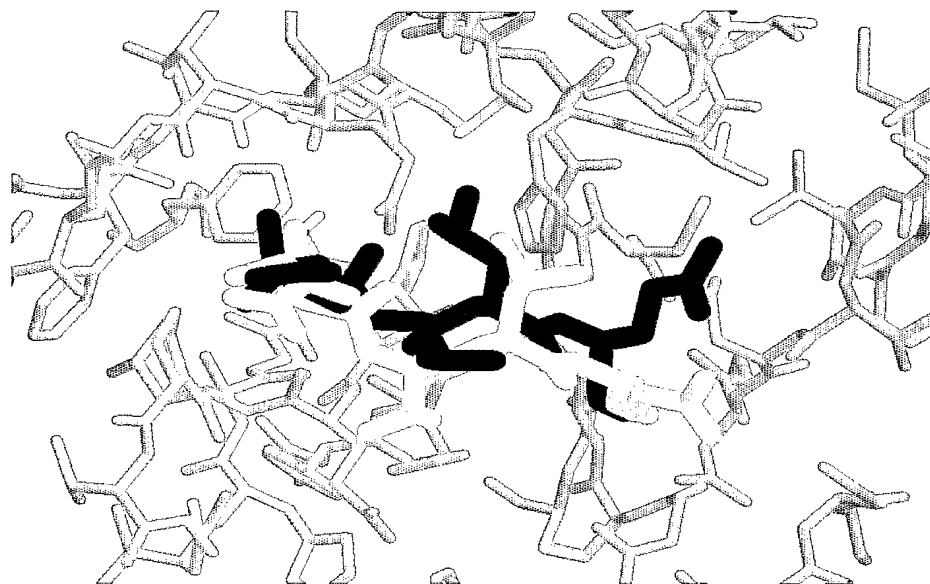
第93図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
CYS145 N	O.co2	100	2.70
MET165 CG	C.3	100	4.00
GLU166 N	O.2	100	2.70
THR190 N	O.3	100	2.70

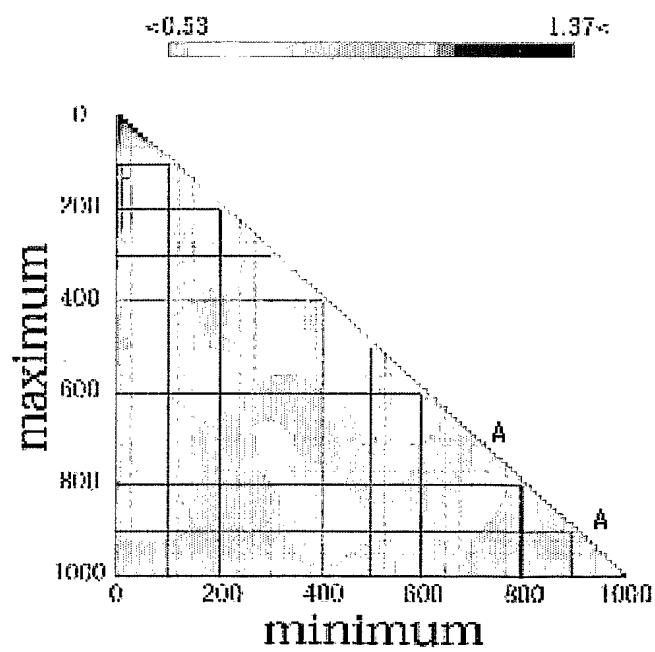
第94図

順位	エネルギー	PDB code	備考
1	-1047.3743	1KZL	transferase
2	-860.437	1J71	hydrolase
3	-844.8737	3UAG	ligase
4	-837.6255	1LKL	complex (tyrosine kinase/peptide)
5	-829.8176	1QF4	ligase
6	-732.2087	1A2N	transferase
7	-721.6213	1G1F	hydrolase, signaling protein
8	-698.5922	1F7B	lyase
9	-689.1472	1BFZ	n-terminal product peptide
10	-646.7943	148L	hydrolase(o-glycosyl)
11	-634.4654	1CGL	metalloprotease
12	-629.1673	1JIL	ligase
13	-616.8733	1FF1	signaling protein
14	-611.1171	1F9E	apoptosis
15	-567.0738	1R1R	oxidoreductase
16	-554.5321	1I95	ribosome
17	-547.2494	1FQX	hydrolase
18	-536.7069	1HCT	complex (signal transduction/peptide)
19	-531.1014	1SIA	mucin motif
20	-508.9899	1JIJ	ligase
21	-507.9655	1LSP	hydrolase (o-glycosyl)
22	-497.6341	1F8H	endocytosis/exocytosis
23	-492.3974	1F74	lyase
24	-443.232	1QH5	hydrolase
25	-427.5925	1JII	ligase
26	-417.4991	1JQY	toxin
27	-416.9956	2KCE	methyltransferase
28	-396.7898	1EJB	transferase
29	-387.6441	1MMJ	hydrolase
30	-358.2162	1SLY	glycosyltransferase
39	-245.9500	1UK4	hydrolase

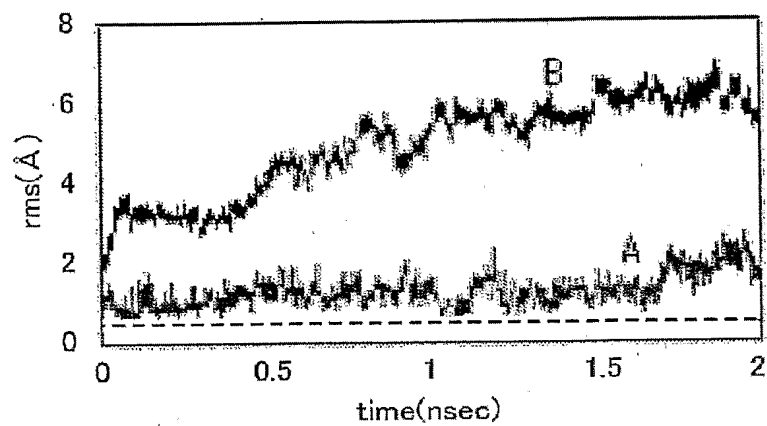
第95図



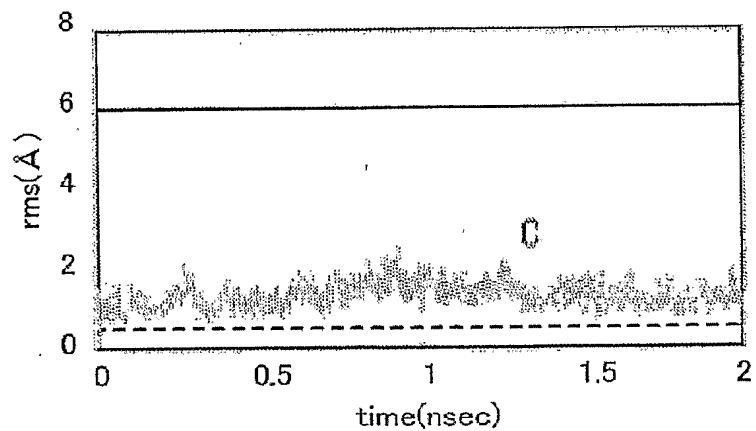
第96図



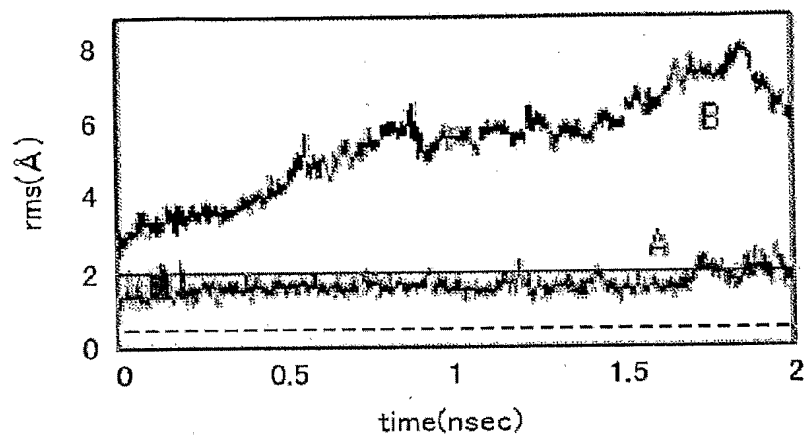
第97図



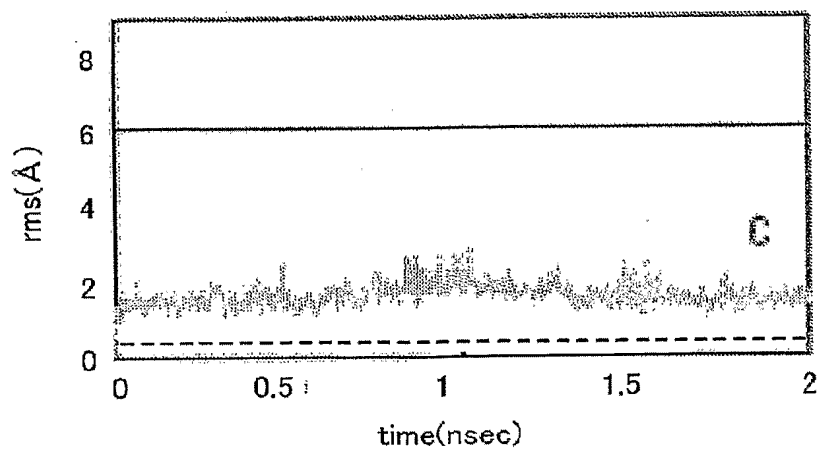
第98図



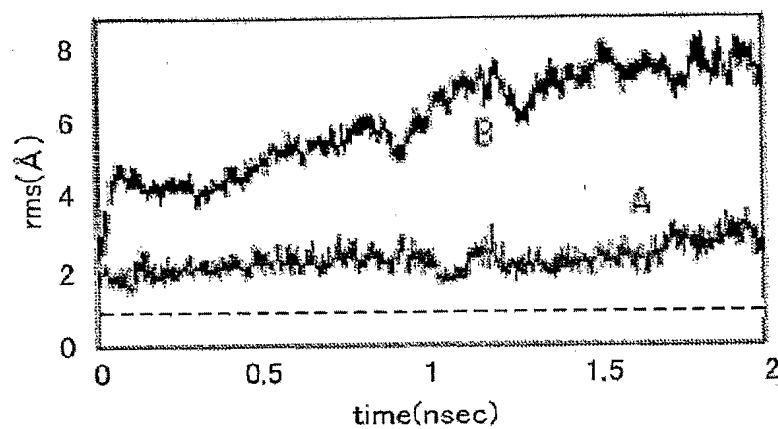
第99図



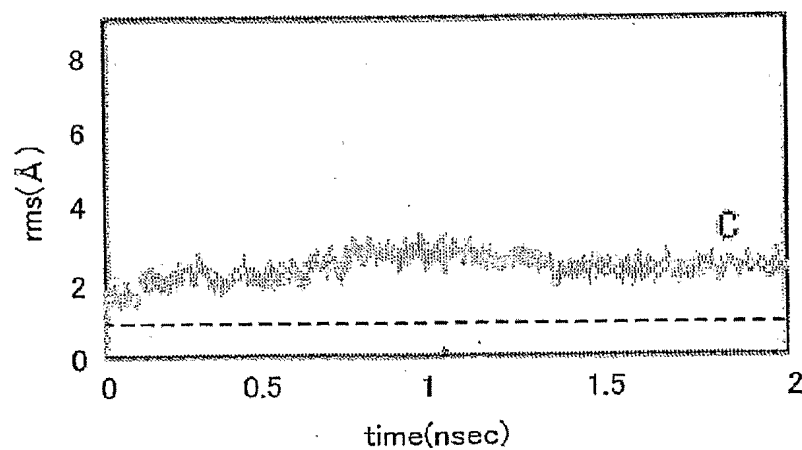
第100図



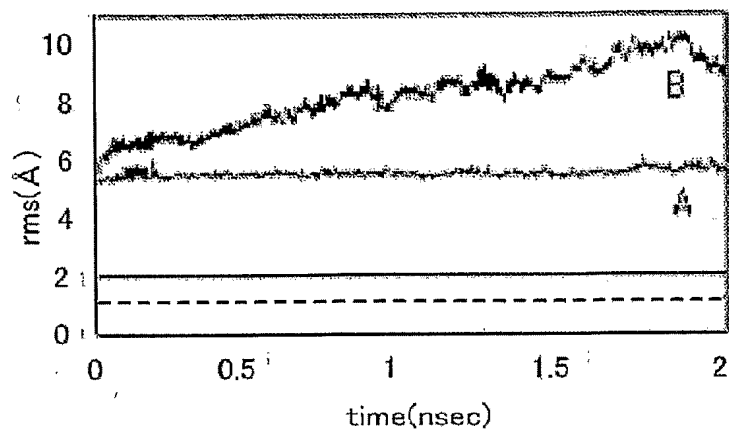
第101図



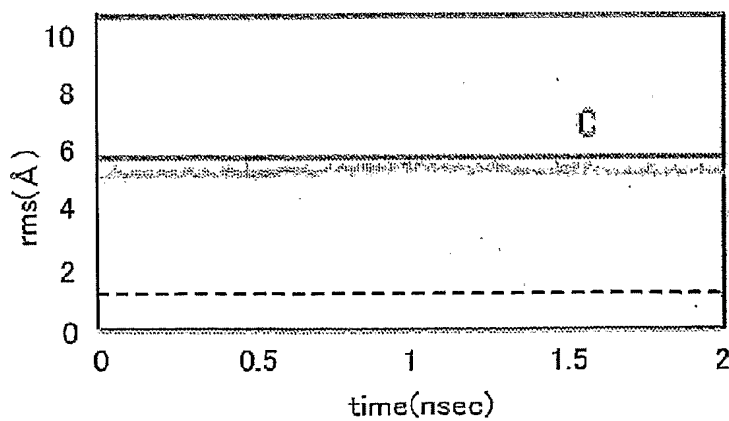
第102図



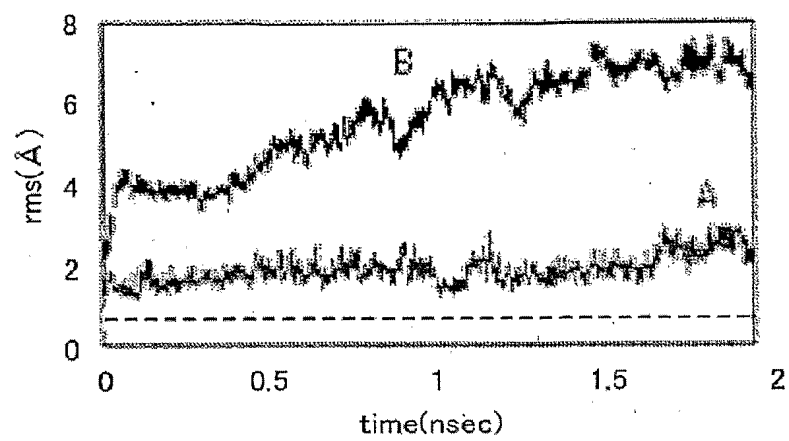
第103図



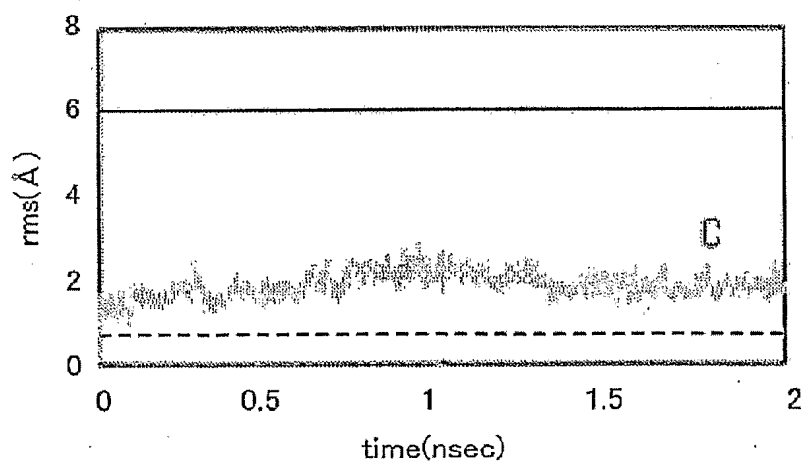
第104図



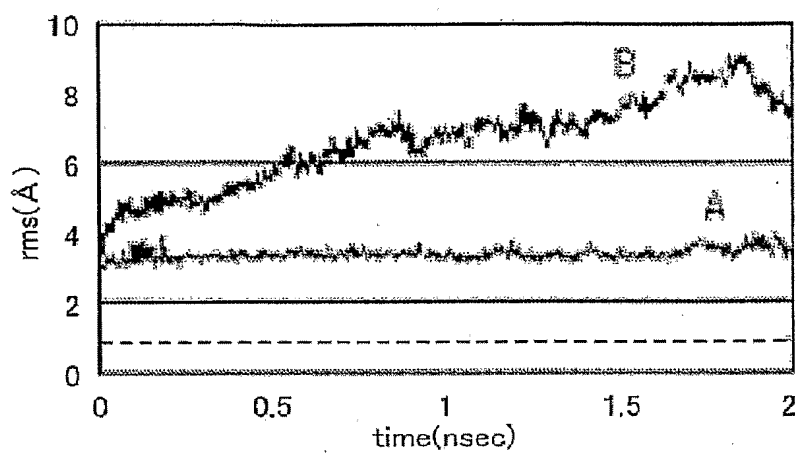
第105図



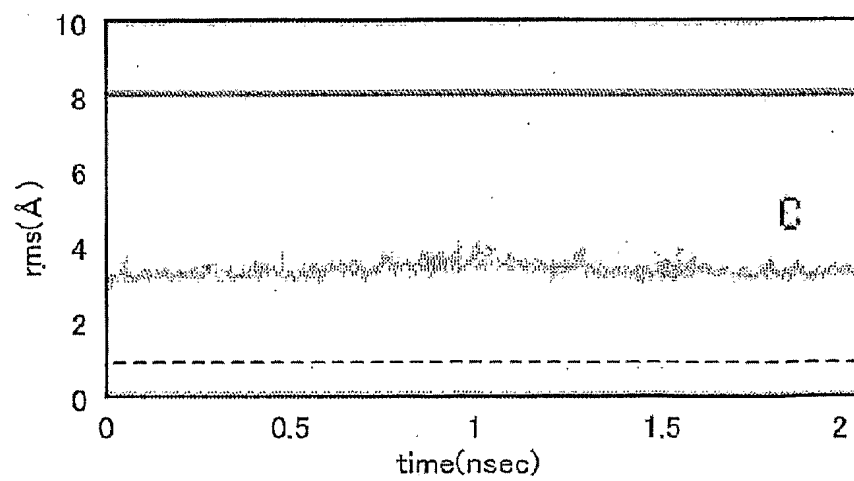
第106図



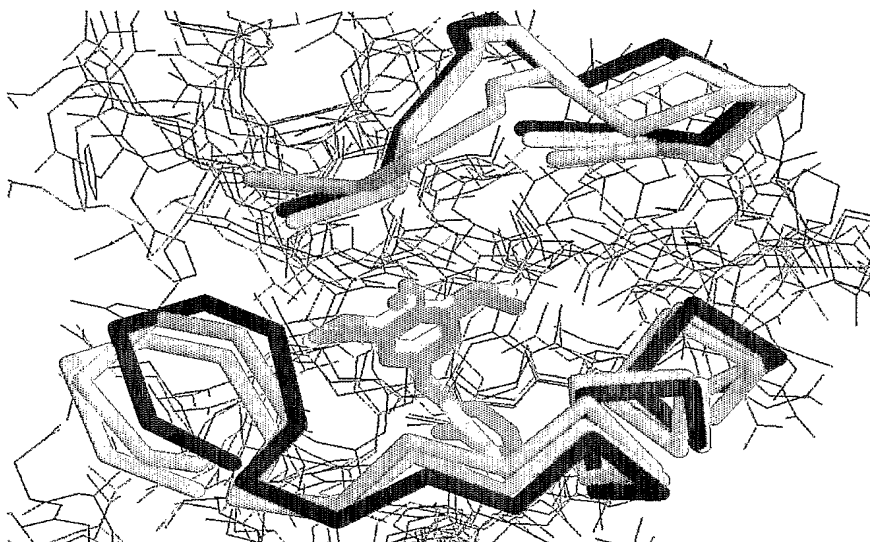
第107図



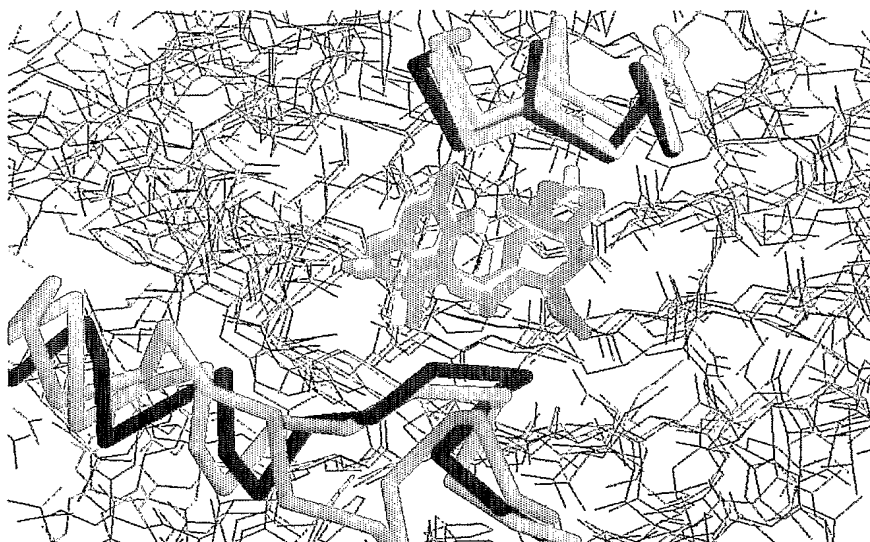
第108図



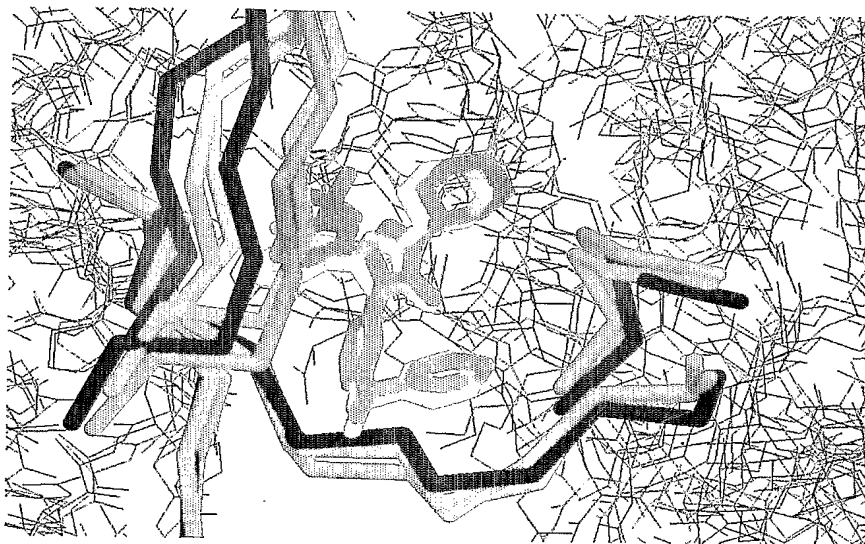
第109図



第110図



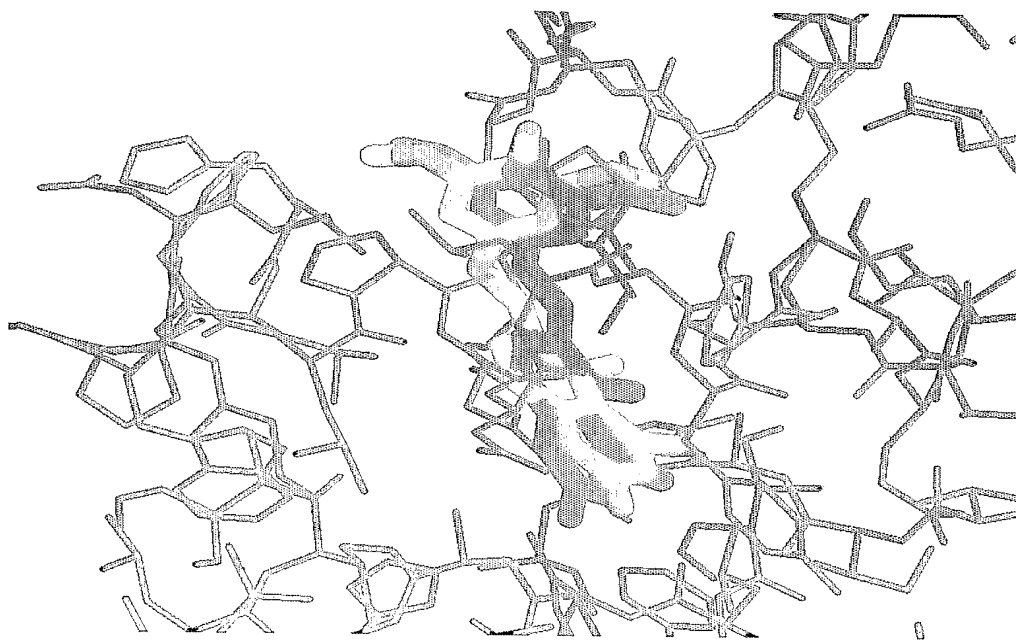
第111図



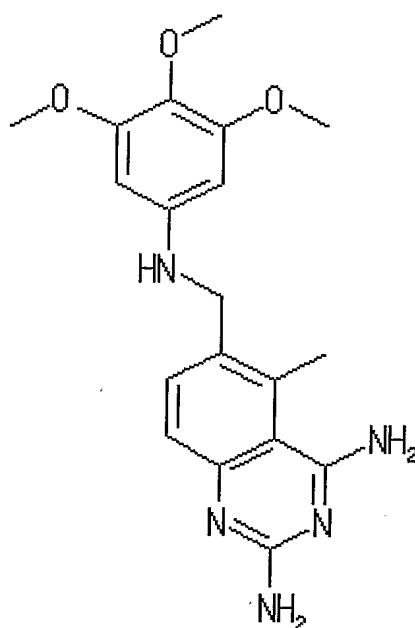
第112図

活性部位の原子	リガンドの原子タイプ	相互作用の強さ	相互作用の距離(Å)
LUE4 O	N.pl3	100	2.87
ASP26 OD1	N.ar	300	3.00
ASP26 OD2	N.pl3	300	3.00

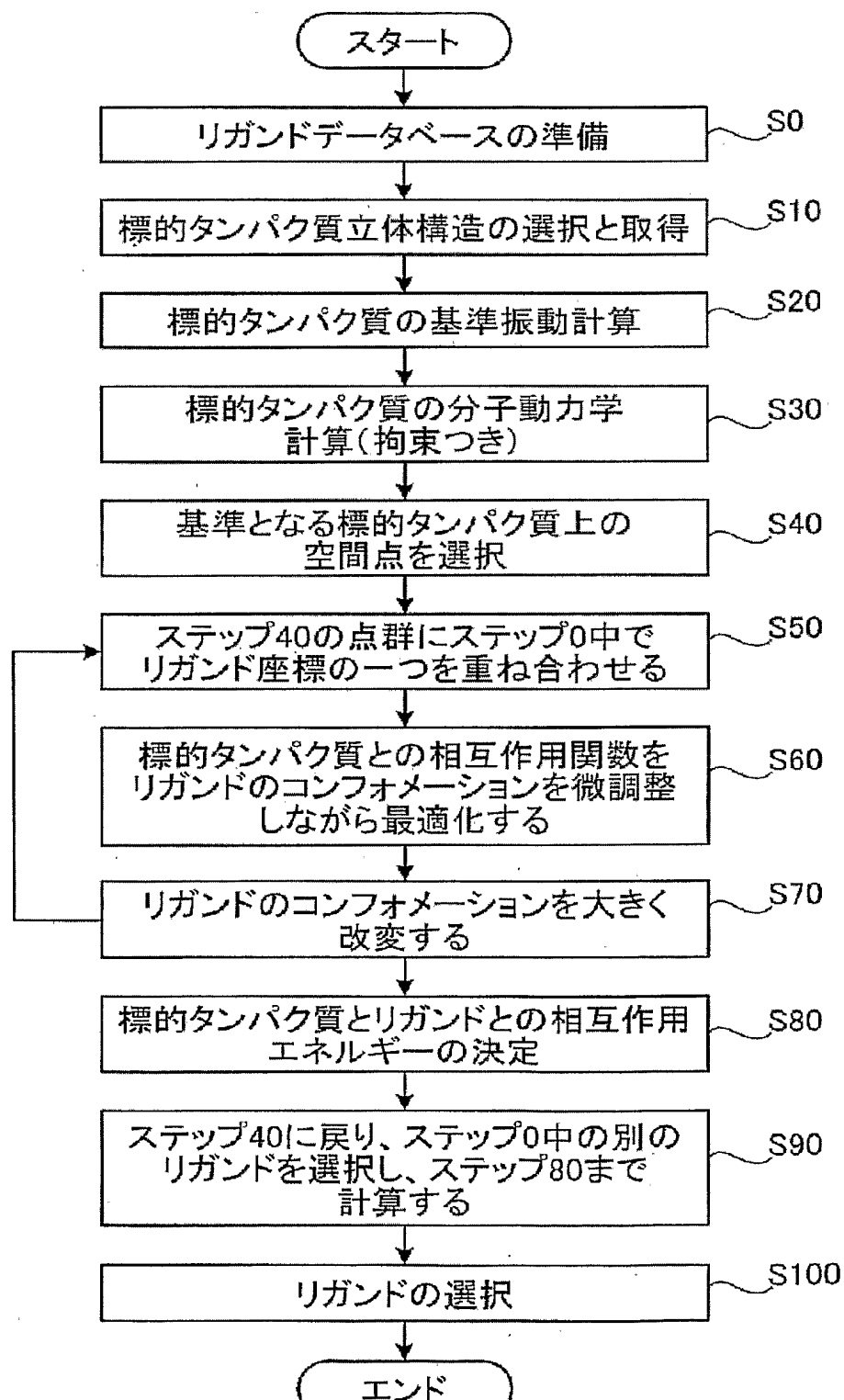
第113図



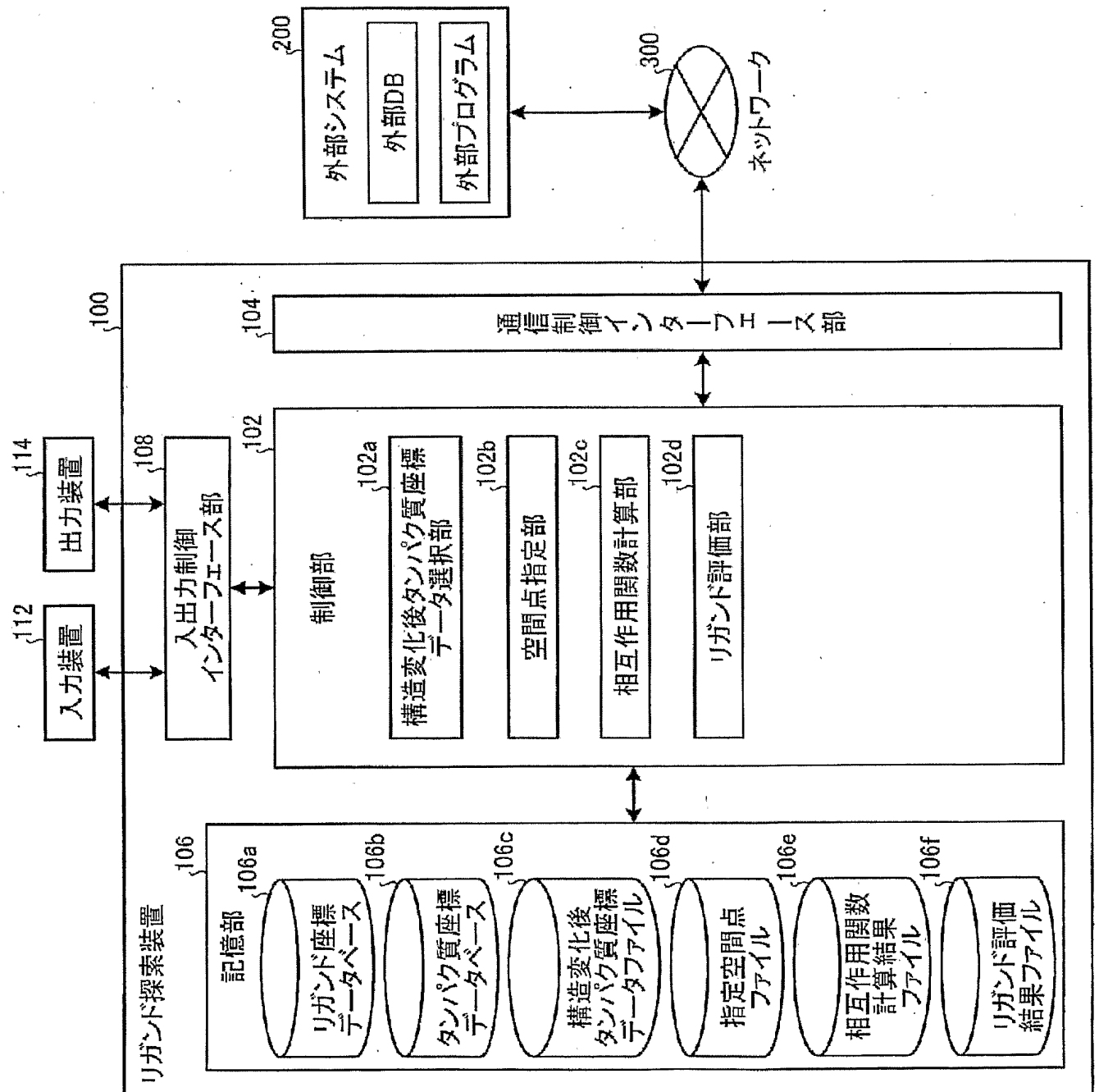
第114図



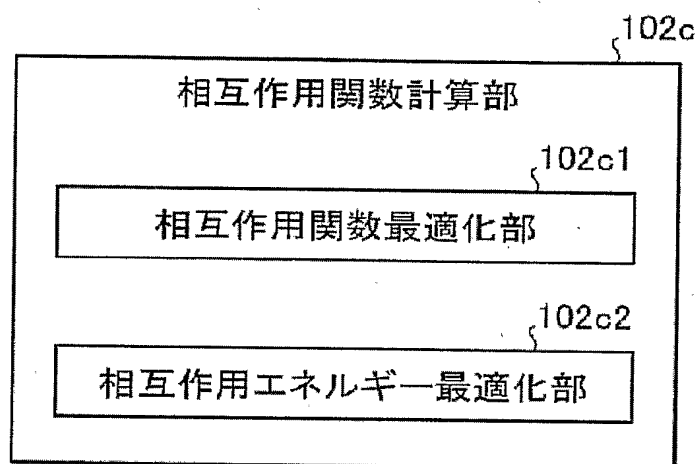
第115図



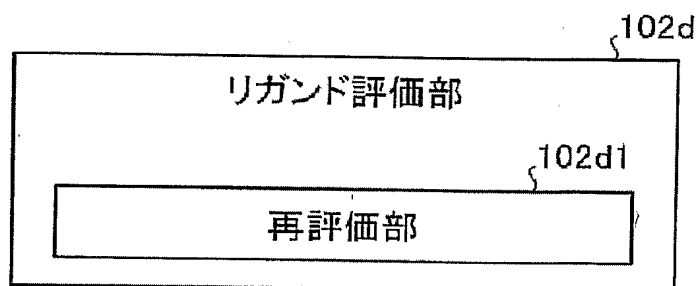
第116図



第117図



第118図



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/003558

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.⁷ G06F19/00, G01N33/48, 33/566, 33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.⁷ G06F19/00, G01N33/48, 33/566, 33/68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2005

Kokai Jitsuyo Shinan Koho 1971-2005 Toroku Jitsuyo Shinan Koho 1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPlus (JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 2002/057954 A1 (Kabushiki Kaisha Inshiriko Saiensu), 25 July, 2002 (25.07.02), Full text; Figs. 1 to 16 (Family: none)	1-5, 8-15, 18-25, 28-31 6-7, 16-17, 26-27
Y A	WO 1996/013785 A1 (Akiko ITAI), 09 May, 1996 (09.05.96), Full text; Figs. 1 to 4 & US 2001/0018682 A1 & EP 0790567 A1	1-5, 8-15, 18-25, 28-31 6-7, 16-17, 26-27
A	WO 1993/020525 A1 (Akiko ITAI), 14 October, 1993 (14.10.93), Full text; Figs. 1 to 8 & US 5642292 A & EP 0633534 A1	1-31



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

24 June, 2005 (24.06.05)

Date of mailing of the international search report

19 July, 2005 (19.07.05)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/003558

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	Hideaki UMEYAMA, "Tanpakushitsu no Yudo Tekigo o Koryoshita Ligand Tansaku System", Dai 32 Kai Kozo Kassei Soka Symposium Koen Yoshishu, 10 November, 2004 (10.11.04), pages 33 to 36	1-31
A	Noriyuki YAMATO, "Hyoteki Tanpakushitsu no Induced Fit o Koryoshita Ligand Docking (1): Brownian Doryoku Gakuho o Mochiita Fukugotai Kozo Saitekika", Dai 31 Kai Kozo Kassei Soka Symposium Koen Yoshishu, 2003, pages 115 to 118	1-31

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl.⁷ G06F19/00, G01N33/48, 33/566, 33/68

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl.⁷ G06F19/00, G01N33/48, 33/566, 33/68

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JSTPlus (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y A	WO 2002/057954 A1 (株式会社インシリコサイエンス) 2002. 07. 25, 全文, 第 1-16 図 (ファミリーなし)	1-5, 8-15, 18-25, 28-31 6-7, 16-17, 26-27
Y A	WO 1996/013785 A1 (板井昭子) 1996. 05. 09, 全文, 第 1-4 図 & US 2001/0018682 A1 & EP 0790567 A1	1-5, 8-15, 18-25, 28-31 6-7, 16-17, 26-27

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の 1 以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

24. 06. 2005

国際調査報告の発送日

19.7.2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
郵便番号 100-8915
東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官 (権限のある職員)

宮久保 博幸

5 L

3 1 3 6

電話番号 03-3581-1101 内線 3562

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	WO 1993/020525 A1 (板井昭子) 1993.10.14, 全文, 第1-8図 & US 5642292 A & EP 0633534 A1	1-31
P, X	梅山秀明, “タンパク質の誘導適合を考慮したリガンド探索システム”, 第32回構造活性相関シンポジウム講演要旨集, 2004.11.10, p.33-36	1-31
A	山乙教之, “標的蛋白質の Induced Fit を考慮したリガンドドッキング (1): ブラウン動力学法を用いた複合体構造最適化”, 第31回構造活性相関シンポジウム講演要旨集, 2003, p.115-118	1-31